

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: RITA MITRA Examiner #: 77775 Date: 11/23/01
 Art Unit: 1653 Phone Number: 301-605-1211 Serial Number: 09/617099
 Mail Box and Bldg/Room Location: 9B01, 2MT Results Format Preferred (circle): PAPER DISK E-MAIL
4203

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Protein RIM2

Inventors (please provide full names): Susumu Seino, Tadao Shibasaki, 1
Nobuaki Ozaki

Earliest Priority Filing Date: 10/3/1999

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

I request a literature search and a patent search on "Protein RIM2".
 - please see attached information on technical field of the invention.
 - Claims 3-12 have been elected and will be examined.

Keywords:

- RIM Protein and isoform
- interacts with Rab3 (protein)
- regulator of Rab3-dependent synaptic vesicle fusion
- interacts with GTP/GDP exchange factor (GEF)
- regulation of exocytosis of neurons and endocrine cells
- recombinant vector
- DNA encoding RIM protein

Point of Contact:
 Mary Hele
 Technical Info. Specialist
 CMM 12D16 Tel: 303-4258

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) _____	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: _____	Bibliographic <input checked="" type="checkbox"/>	Dr. Link _____
Date Completed: <u>11/30/01</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems _____
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____

THIS PAGE BLANK (USPTO)

See ID NO. 2
AC. NO. AAV40485
Database: N-gene-seq-0601

Oligonucleotide D1
Oligonucleotide D1
Oligonucleotide D1
Oligonucleotide D2
Oligonucleotide D2
Oligonucleotide D1
Human Doc2-beta gene
Gene encoding a su
Sequence encoding
Human protein kinase
DNA encoding a pro
Human low adenosine
Human adenosine re
Human low adenosine
Human adenosine re
Murine LOBO homo
An EcoRI fragment
Kaposi's sarcoma-a
KSHV LTR DNA (nucl
KSHV long unique c
polyglutamine trac
Mouse Doccalpha ge
plasmid/falcipla
Exons E, C and A O
Exons D, C, B and
Spino cerebellar at
Spondylosis insolut pol
Mouse sequence of hu
DNA sequence of hu
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Nucleotide rich reg

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

3M nucleic acid - nucleic search, using sw model

Run on: November 21, 2001, 18:56:16 ; Search time 358.58 Seconds
(without alignments)
8720.369 Million cell updates/sec

Title: US-09-617-099b-2

Sequence: 1 gctccctaggtgtgttcg.....acattgtccctagcaag 4980

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_0601:*

1: /cgnl_9/gcgdata/geneseq/geneseqn/NA1980.DAT:*

2: /cgnl_9/gcgdata/geneseq/geneseqn/NA1981.DAT:*

3: /cgnl_9/gcgdata/geneseq/geneseqn/NA1982.DAT:*

4: /cgnl_9/gcgdata/geneseq/geneseqn/NA1983.DAT:*

5: /cgnl_9/gcgdata/geneseq/geneseqn/NA1984.DAT:*

6: /cgnl_9/gcgdata/geneseq/geneseqn/NA1985.DAT:*

7: /cgnl_9/gcgdata/geneseq/geneseqn/NA1986.DAT:*

8: /cgnl_9/gcgdata/geneseq/geneseqn/NA1987.DAT:*

9: /cgnl_9/gcgdata/geneseq/geneseqn/NA1988.DAT:*

10: /cgnl_9/gcgdata/geneseq/geneseqn/NA1989.DAT:*

11: /cgnl_9/gcgdata/geneseq/geneseqn/NA1990.DAT:*

12: /cgnl_9/gcgdata/geneseq/geneseqn/NA1991.DAT:*

13: /cgnl_9/gcgdata/geneseq/geneseqn/NA1992.DAT:*

14: /cgnl_9/gcgdata/geneseq/geneseqn/NA1993.DAT:*

15: /cgnl_9/gcgdata/geneseq/geneseqn/NA1994.DAT:*

16: /cgnl_9/gcgdata/geneseq/geneseqn/NA1995.DAT:*

17: /cgnl_9/gcgdata/geneseq/geneseqn/NA1996.DAT:*

18: /cgnl_9/gcgdata/geneseq/geneseqn/NA1997.DAT:*

19: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT:*

20: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:*

21: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:*

22: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	844.8	17.0	1302	19	AAV40485
2	377	7.6	2276	21	AAFC6011
3	207.6	4.2	355	20	AAV90372
4	207.6	4.2	355	20	AAV90372
5	195.6	3.9	378	21	AACT4934
6	93.4	1.9	936	22	AAFS8252
7	93.4	1.9	936	22	AAFS8254
8	93.4	1.9	936	22	AAFS8257
9	93.4	1.9	936	22	AAFS8259
10	93.4	1.9	936	22	AAFS8262
11	93.4	1.9	938	22	AAFS8255

C 12	92.8	1.9	936	22	AAFS8252	Oligonucleotide D1
C 13	92.8	1.9	936	22	AAFS8254	Oligonucleotide D1
C 14	92.8	1.9	936	22	AAFS8257	Oligonucleotide D1
C 15	92.8	1.9	936	22	AAFS8259	Oligonucleotide D2
C 16	92.8	1.9	936	22	AAFS8262	Oligonucleotide D2
C 17	92.8	1.9	936	22	AAFS8265	Oligonucleotide D1
C 18	55.4	1.1	2043	18	AAAT79627	Human Doc2-beta ge
C 19	50	1.0	10732	21	AAAI0594	Gene encoding a su
C 20	47.6	1.0	2182	9	AAAB0338	Sequence encoding
C 21	47.6	1.0	2244	21	AAZ4431	Human protein kina
C 22	47.6	1.0	2245	16	AAQ79892	DNA encoding a pro
C 23	47.6	1.0	2245	21	AAAF21404	Human low adenosin
C 24	47.6	1.0	2245	21	AAAF35282	Human adenosine re
C 25	47.6	1.0	38644	21	AAAF21424	Human low adenosin
C 26	47.6	1.0	38644	21	AAAF35302	Human adenosine re
C 27	46.6	0.9	49999	20	AAZ23895	Murine LOBO homolo
C 28	45.4	0.9	1159	21	AAAS9240	An EcorI fragment
C 29	45	0.9	3489	21	AAAS9240	Kaposi's sarcoma-a
C 30	45	0.9	32207	20	AAV73805	KSHV LUR DNA (nuc
C 31	45	0.9	137507	19	AAV19941	Polyglutamine tranc
C 32	44.4	0.9	486	22	AAFT5507	Mouse Doc2alpha ge
C 33	44.4	0.9	2255	20	AAZ00321	Plasmodium falcipa
C 34	44.4	0.9	3579	21	AAAF0099	Exons E, C and A o
C 35	43.8	0.9	1037	21	AAAS9242	Exons D, C, B and
C 36	43.8	0.9	1472	21	AAAS9241	Spinocherebellar at
C 37	43.4	0.9	397	20	AAAB9891	Mouse inositol pol
C 38	43.2	0.9	1876	17	AAAT29743	DNA sequence of hu
C 39	43	0.9	5120	22	AAAC84677	Nucleotide sequenc
C 40	42.4	0.9	4466	21	AAAI4663	Nucleotide sequenc
C 41	42.4	0.9	4478	21	AAAI4661	Nucleotide sequenc
C 42	42.4	0.9	4547	21	AAAI4664	Nucleotide sequenc
C 43	42.4	0.9	4571	21	AAAI4662	Nucleotide sequenc
C 44	42.4	0.9	77536	21	AAAI4651	Nucleotide sequenc
C 45	42	0.8	203	19	AAV30271	Glutamine rich reg

ALIGNMENTS

RESULT 1	
AAV40485	standard; cDNA: 1302 BP.
ID	AAV40485;
AC	AAV40485;
XX	
XX	09-NOV-1998 (first entry)
DE	Human secreted protein CO618_1 cDNA.
XX	
KW	Secreted protein: CO618_1; human; ds.
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS 123..1247
XX	/*tag= a
XX	
XX	WO9831802-A1.
XX	
XX	23-JUL-1998.
XX	
XX	21-JAN-1998; 98WO-US01007.
XX	
XX	20-JAN-1998; 98US-0010047.
XX	21-JAN-1997; 97US-0072103.
XX	16-JUN-1997; 97US-0877035.
XX	
XX	(GENY) GENETICS INST INC.
XX	
XX	Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
XX	Racie LA, Spaulding V, Treacy M;
XX	WPI: 1998-414101/35.

SEE ID NO: 2
AC. NO. AAV40485
Database: N-Geneseq-0601

Homo sapiens.
W0200058356-A1.

XX 05-OCT-2000.
 PD 22-MAR-2000; 2000WO-US07535.
 PF 26-MAR-1999; 99US-0126511.
 PR 17-DEC-1999; 99US-0172413.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM, Komatsoulis G;
 PI WPI: 2000-594639/56.
 DR P-PSDB; AAB34818.
 XX Fifty nucleic acid molecules encoding human secreted proteins, useful
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases -
 XX
 PS Claim 1; Page 372-373; 425pp; English.
 XX The polynucleotide sequences given in AAC59966 to AAC60015 encode the
 CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
 CC AAB34852 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities include:
 CC cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 CC vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for
 CC preventing, treating or ameliorating a medical condition in e.g. humans,
 CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities. The polynucleotides are
 CC useful for chromosome identification. They are also useful as probes for
 CC diagnosing a disorder related to the female reproductive system,
 CC particularly breast and/or ovary cancer. They are also useful in the gene
 CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
 CC agonists and antagonists from the present invention are useful in the
 CC diagnosis, treatment and prevention of cancer, immune disorders, and
 CC cardiovascular disorders, wound healing, neurological diseases and
 CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
 CC used in the exemplification of the present invention.
 XX
 XX Sequence 2276 BP; 699 A; 451 C; 436 G; 689 T; 1 other;

Query Match 7.6%; Score 377; DB 21; Length 2276;
 Best Local Similarity 71.9%; Pred. No. 2.6e-97;
 Matches 507; Conservative 0; Mismatches 195; Indels 3; Gaps 1;

QY 4100 aaaaagctacgagcactgttcagagaagacgagacccggctcagtcagtgagatgag 4159
 DB 96 aaaaagttaaaagcattccacagagaagacagacagcagtcagtgagatgaga 155
 QY 4160 aacgtgagaccgcccagccgagcagatccacagatggcagctgaacacactaagc 4219
 DB 156 aa---gattggtaaagccagccgagcagtgctactgattggtccagatcaacattaaagc 212
 QY 4220 tcggaagaatctcatcttcctgggtgcgcgcctgtcctcagccagcagttcagatg 4279
 DB 213 tctgagggcaatttaatttctctgagtgagcagctggagagctgcaggtcaattcagtgat 272
 QY 4280 tctctgagtgagcctgggcccctgtctcagtagtggagccagacccctgtactcctgca 4339
 DB 273 tttctgtgattgtgggacccagcagctgtgtggcccaaaccttgcacccctgca 332
 QY 4340 atgggtgacattcagtggtgagatgataaaagggcagcagctgtagatgaatcattc 4399
 DB 333 atgggtgataatacaaatgagtagacaaaagggcagcagctgtagatgaatgagatcatt 392
 QY 4400 cgggcgcgcgcctgtgtgtaaaccaaggttcacaagacactgccagcagcagatgtcaag 4459

DB 393 agagacgaagcctcacaacaagaagcctgttccaactctacacactgtccatatgtcaaa 452
 QY 4460 gttatctgttagacaacgagatctgcatagcacaagaagaacgaagtcgcagaag 4519
 DB 453 gtatacttttggaaatggygcctgtatagccaagaagaagaatgtgcagcaaa 512
 QY 4520 accctggagccctgtaccagaagcctctgtccttcgtgaggaagcccccggaggaggtg 4579
 DB 513 acccttgatccttgatcagaagcctctgtgttttgatgaaagtcacagggtaagtt 572
 QY 4580 ttacagatcatgtcttggggaatataggcgtatgatcacaaactcttatggagatg 4639
 DB 573 cttaagtgattgtctggggagactatgacagatgaccacaatgtcttatgggtgtg 632
 QY 4640 gccacgaactctttagatgaactgtgaactacatacaatgtgattgtgattcaactc 4699
 DB 633 gtcagactcttgtggaagaactgcacctgtccaagatgtagtcgattgtacaaattg 692
 QY 4700 ttcctctcctcctccttagatcacaactcgcacactcgtgacaagaagacttccaa 4759
 DB 693 ttccacacgtctcctcactgtgtgataccacaactacccctcaccggcggttccag 752
 QY 4760 tgcctctggaaagttctaccggacctcttactcgttcataag 4804
 DB 753 tcatctcgaaagttcaactgtgacctcctcctgtatcgtacataag 797

RESULT 3
 AAV90372/c
 ID AAV90372 standard; cDNA; 355 BP.
 XX
 AC AAV90372;
 XX
 XX 15-FEB-1999 (first entry)
 DT
 XX
 DE EST clone DM18.
 XX
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokines; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 OS
 XX
 PN W09845436-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 10-APR-1998; 98WO-US06955.
 XX
 PR 10-APR-1997; 97US-0838821.
 PA (GENY) GENETICS INSTR INC.
 XX
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 DR WPI: 1999-070077/06.
 XX
 XX New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.
 XX
 PS Claim 1; Page 519-520; 618pp; English.
 CC The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating

CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC adherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 XX
 SS Sequence 355 BP; 93 A; 82 C; 83 G; 97 T; 0 other;

Query Match 4.2%; Score 207.6; DB 20; Length 355;
 Best Local Similarity 90.2%; Pred. No. 2.9e-49;
 Matches 222; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 3844 cagtgcttaccatccaatgcaaaacagacagatggcggtgagggaagactgac 3903
 Db 247 CAGTGCTTTACATCCAAATGCAAGACAGACAAATGGCCATATCAGGAAGAACATGAC 188
 QY 3904 caaagaccagcagatcagtgagacagatgctcactggaagaatgacggcagcagtc 3963
 Db 187 AAAAGCACCACGATCATGATGAGACATGCTCACTGAGAGATGATGACACCACTC 128
 QY 3964 cgacactgcagtgaggcgccctgggtacagtgagcaagaagcgagatcagcaltgggc 4023
 Db 127 TGACATCGCAGTGGGCGACCTTGGCCACCAAGGCGGCTTAAGCCCTTGGTGC 68
 QY 4024 caaatgtagctatgttgctctcaccgaaaagtcgacgtgacctcaactcagca 4083
 Db 67 CAAATGTGACGTAATGCTGTGCTGCTACGAAAGCGCAGTGTCTCAGCTCAACGG 8
 QY 4084 aaccca 4089
 Db 7 ACTCGA 2

RESULT 4
 ID AAV89630
 XX AAV89630 standard; cDNA; 355 BP.

AC AAV89630;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE EST clone CO618.
 XX

Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.

XX Homo sapiens.
 OS

XX WO9845436-A2.
 PN

XX 15-OCT-1998.
 PD

XX 10-APR-1998; 98WO-US06955.
 PF

XX 10-APR-1997; 97US-0838821.
 PR

XX (GEMV) GENETICS INST INC.
 PA

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie JA, Spaulding V, Treacy M;
 DR WPI; 1999-070077/06.

XX New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.
 XX
 PS Claim 1; Page 279; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST).
 CC

CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 XX
 SS Sequence 355 BP; 97 A; 83 C; 82 G; 93 T; 0 other;

Query Match 4.2%; Score 207.6; DB 20; Length 355;
 Best Local Similarity 90.2%; Pred. No. 2.9e-49;
 Matches 222; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 3844 cagtgcttaccatccaatgcaaaacagacagatggcggtgagggaagactgac 3903
 Db 109 cagtgcttaccatccaatgcaaaacagacagatggcggtgagggaagactgac 168
 QY 3904 caaagaccagcagatcagtgagacatgctcactggggaagaatgacggcagcagtc 3963
 Db 169 aaaaagaccagcagatcagtgagacatgctcactggggaagaatgacggcagcagtc 228
 QY 3964 cgacactgcagtgaggcgccctgggtaccagtgcaagaagcgcgatcagcaltgggc 4023
 Db 229 tgacactgcagtgaggcgccctgggtaccagtgcaagaagcgcgatcagcaltgggc 288
 QY 4024 caaatgtagctatgttgctctcaccgaaaagtcgacgtgacctcaactcagca 4083
 Db 289 caaatgtagctatgttgctctcaccgaaaagtcgacgtgacctcaactcagca 348
 QY 4084 aaccca 4089
 Db 349 actcga 354

RESULT 5
 ID AAC74934
 XX AAC74934 standard; cDNA; 378 BP.

AC AAC74934;
 XX

DT 08-FEB-2001 (first entry)
 XX

XX Human ORFX ORF489 polynucleotide sequence SEQ ID NO:977.
 DE

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antiproscler; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotonic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; anti-inflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; anti-inflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.
 OS

XX WO200058473-A2.
 PN

XX 05-OCT-2000.
 PD

XX 31-MAR-2000; 2000WO-US08621.
 PF
 XX

2R 31-MAR-1999; 99US-0127607.
 CC 02-APR-1999; 99US-0127636.
 2R 05-APR-1999; 99US-0127728.
 2R 30-MAR-2000; 2000US-0540763.
 XX
 2A (CURA-) CURAGEN CORP.
 XX
 2I Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB40725.
 XX
 2T Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 5; Page 956; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiallergic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatologic; antitumor; vasodilator;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease: to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 378 BP; 85 A; 106 C; 130 G; 57 T; 0 other;

Query Match 3.9%; Score 195.6; DB 21; Length 378;
 Best Local Similarity 71.3%; Pred. No. 8.2e-46;
 Matches 258; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 4119 ttccagagaacacagagacgagctagctgagatgaggaactgtatgacccagcagg 4178
 Db 17 tccagagagacagcagagagcgctgagctgagatgagatgcccgccgacactgcccagcagg 76
 QY 4179 ccagcgggaatccacagatgagcagcatgaacagctatagctcggaagaaatctgact 4238
 Db 77 ccagcagagatgcatctgagcagcatgaacagctatgagcagggggaaccttaact 136
 QY 4239 tccctgggtccgctgctgctgacgacgactgagttgattcccgatgagcctggcc 4288
 Db 137 atgagagagcttgcctgctgctgacgacgactgagttgattcccgatgagcctggcc 196
 QY 4299 ctgtcacaactagtgagacgacgacgacgactgactcctgcaatgggtgacatcgaatg 4358
 Db 197 cggcacaactgtgtggcgccgacgacgactgacccacacccatggggatgtgagatcg 256
 QY 4359 gaatgatgataaagaagacagctgtgagtagaatacatccggcgcgcgctgttg 4418
 Db 257 gtctgcagagcagcagcagctgagttgagatgacattacagctcggtgagcag 316
 QY 4419 taacaacaggttccaaagacacgacgacgacgacgacgacgacgacgacgacgacgacg 4478
 Db 317 ccaagcagagcctccaaagacacgacgacgacgacgacgacgacgacgacgacgacgacg 376
 QY 4479 ga 4480

Db 377 ca 378

RESULT 6

ID AAF58252 standard; DNA; 936 BP.

XX AAF58252;

DT 24-APR-2001 (first entry)

DE Oligonucleotide D1835.

KM Electron-transfer group; ETM; mismatch; genotyping;

KW gene expression; ss.

OS Synthetic.

PN W0200107665-A2.

PD 01-FEB-2001.

PF 26-JUL-2000; 2000WO-US20476.

PR 26-JUL-1999; 99US-0145695.

PR 17-MAR-2000; 2000US-0190259.

PA (CLIN-) CLINICAL MICRO SENSORS INC.

PI Umek RM;

DR WPI; 2001-159728/16.

PT Nucleic acids containing electron-transfer group, useful as labels in

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on

PT a single surface -

PS Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic

CC acids each containing an electron-transfer group (ETM) having

CC different redox potentials. The invention is used for electronic

CC detection of nucleic acids, especially of substitutions (mismatches)

CC and single-nucleotide polymorphisms, e.g. for genotyping,

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

CC monitoring gene expression.

```
QY 2399 aaatggaccgtcctctatataccgttaccccaacgagtcgtgagatgagat 2458
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 342 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 401
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2459 gtcccgagttctatctgtgacagcttcaataaactatggttgacaaggtgtgcac 2518
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 461
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2519 cagttgattgatacttttggagcaaatgctccctccagaggaagtggagccca 2578
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 462 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2579 aggaatccttatgttaagatttactcctccacagataagtgataaataagaaga 2638
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 522 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 581
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2639 acaaaaacagtcagaacttttggaaacccaatggaacgacttcaattatctcct 2698
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 641
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2699 gtccacggaagaattcgtgaaacgaatgctggaattacccttgggacaaactaga 2758
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 642 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 701
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2759 gtccgagaagaagagcgaattctctagagagaatttaattgaattgaaacagcttg 2818
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 702 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 761
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2819 ctgagatgtagcgccactgtataagctgacagcccatgagtctcctcattgcac 2878
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 762 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 821
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2879 cctcgccctcccaatctgcccc 2903
      || || || || || || || || || || || || || || || || || || || || ||
Db 822 cccccccccccccccccccccccc 846

RESULT 7
AAF58254
ID AAF58254 standard; DNA; 936 BP.
AC AAF58254;
XX
XX 24-APR-2001 (first entry)
DE Oligonucleotide D1875.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PE 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI: 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PS a single surface
XX
XX Example 6; Page 127; 159pp; English.
```

CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 1.9%; Score 93.4; DB 22; Length 936;
Best local similarity 4.2%; Pred. No. 2,8e-16;
Matches 34; Conservative 435; Mismatches 336; Indels 0; Gaps 0;

```
QY 2099 ttaaggattgtaggaagaatgactgaatcagtcgacttgtgcatattaccaca 2158
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 101
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2159 gtaaaaaagggaatttagctgactgttagacalcctagaccaggtgatgacttg 2218
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2219 gaatggaatgggaggtatgtcaaggagccaatltgaggaaatttacaattatctc 2278
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 221
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2279 gaatccaaacctgaaccacaagltgagctgtgttccaaggccaattggagatctc 2338
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2339 agaataccgtatagcagcagatgcacaacgtaatccagttcgtccattgatactcaa 2398
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 282 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 341
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2399 aaatggaccgtcctctatataccgttaccccaacgagtcgtgagatgagat 2458
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 342 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 401
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2459 gtcccgagttctatctgtgacagcttcaataaactatggttgacaaggtgtgcac 2518
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 461
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2519 cagttgattgatacttttggagcaaatgctccctccagaggaagtggagccca 2578
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 462 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2579 aggaatccttatgttaagatttactcctccacagataagtgataaataagaaga 2638
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 522 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 581
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2639 acaaaaacagtcagaacttttggaaacccaatggaacgacttcaattatctcct 2698
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 641
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2699 gtccacggaagaattcgtgaaacgaatgctggaattacccttgggataagctaga 2758
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 642 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 701
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2759 gtccgagaagaagagcgaattctctagagagaatttaattgaattgaaacagcttg 2818
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 702 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 761
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2819 ctgagatgtagcgccactgtataagctgacagcccatgagtctcctcattgcac 2878
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 762 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 821
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2879 cctcgccctcccaatctgcccc 2903
      || || || || || || || || || || || || || || || || || || || || ||
Db 822 cccccccccccccccccccccccc 846

RESULT 8
AAF58257
```


ID	AAF58257	standard; DNA; 936 BP.
xx		
AC	AAF58257;	
xx		
DT	24-APR-2001	(first entry)
xx		
DE	Oligonucleotide D1954.	
xx		
KW	Electron-transfer group; ETM; mismatch; genotyping;	
KM	gene expression; ss.	
xx		
OS	Synthetic.	
xx		
PN	MO200107665-A2.	
xx		
PD	01-FEB-2001.	
xx		
PF	26-JUL-2000; 2000WO-US20476.	
xx		
PR	26-JUL-1999; 99US-0145695.	
PR	17-MAR-2000; 2000US-0190259.	
xx		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.	
xx		
PI	Unek RM;	
DR	WPI; 2001-159728/16.	
xx		
PT	Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface	
xx		
PS	Example 6; Page 127; 159pp; English.	
xx		
CC	The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.	
SO	Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;	
<hr/>		
Query Match	1.9%; Score 93.4; DB 22; Length 936;	
Best Local Similarity	4.2%; Pred. No. 2.8e-16;	
Matches 34; Conservative 435; Mismatches 336; Indels 0; Gaps 0;		
QY 2099	ttaaaggtttagagaagaatgactgaatcaggctgcagtgtgtgcattttaccocaa	2158
Db 42	www.....	101
QY 2159	gtlaaaaaaggaagttagctgatactctgttagacaccttaagccagtgatgaagtcctg	2218
Db 102	www.....	161
QY 2219	gaatggaaatgggagcgctacttcgaaggaagccacatttgaaggaattacaacatttta	2278
Db 162	www.....	221
QY 2279	gaatccaacctgaacacaaagttagacttgttgcgttaaacgaacttggagatatcct	2338
Db 222	www.....	281
QY 2339	agaatacctgtatagacgatcatgcaaacactgtaattccagttcagtcactttgaactcaa	2398
Db 282	www.....	341
QY 2399	aaaatggacgcgttccttctataatccgttacccatcgaagtcctgcacatgcttgaaggat	2458
Db 342	www.....	401
QY 2459	gtcccgaagttcttatctgtgacagacttcaataaactatggttgaacaagtttgytaac	2518

[illegible]

RESULT	9
AAFS58259	
ID	AAFS58259 standard; DNA; 936 BP.
XX	
AC	AAFS58259;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Oligonucleotide D2004.
XX	
KW	Electron-transfer group; ETM; mismatch; genotyping;
KW	gene expression; ss.
XX	
OS	Synthetic.
XX	
PN	WO200107665-A2.
XX	
PD	01-FEB-2001.
XX	
PE	26-JUL-2000; 2000WO-US20476.
XX	
PR	26-JUL-1999; 99US-0145695.
PR	17-MAR-2000; 2000US-0190259.
XX	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.
XX	
PI	Unek RM;
XX	
DR	WPI; 2001-159728/16.
XX	
PT	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT	a single surface -
XX	
PS	Example 6; Page 128; 159p; English.
XX	
CC	The present invention relates to a composition comprising two nucleic
CC	acids each containing an electron-transfer group (ETM) having
CC	different redox potentials. The invention is used for electronic
CC	detection of nucleic acids, especially of substitutions (mismatches)
CC	and single-nucleotide polymorphisms, e.g. for genotyping,

CC monitoring gene expression.

SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match	1.98;	Score 93.4;	DB 22;	Length 936;
-------------	-------	-------------	--------	-------------

Best Local Similarity 4.28; Pred. No. 2.8e-16;

QY	2099	ttaaagttgtgaggaagaagatgactgatacgaagtcgacttgttgcaattattaccaa	2158
Db	42	#####	101
QY	2159	gtaaaaaagaagtttagctgtactctgtagacatcttagaccagtgatgaagtc	2218
Db	102	#####	161
QY	2219	gaatgtgaatggagagcgtatgcaagagcacaattgaggaattacaacattctc	2278
Db	162	#####	221
QY	2279	gaatccaaacctgaaccacaagttgagctgtgtgttcaagccaatgagatattcct	2338
Db	222	#####	281
QY	2339	agaatacctgatagcagcatgcaacactggaatccagttcagtcattgtaactca	2398
Db	282	#####	341
QY	2399	aaatgaccgtctcttatactcgttacatccacatgacctgcgcatgtgagagat	2458
Db	342	#####	401
QY	2459	gtcccgagttctatctgtagacagcttcaataaactgtgtgacaaggtgtgac	2518
Db	402	#####	461
QY	2519	cagttgtagtacaattttggaggaagaagatctccctcgaaggagaatggagcgca	2578
Db	462	#####	521
QY	2579	aggatctctatgttaagattactcttcacagatagaagtgataaataagaaga	2638
Db	522	#####	581
QY	2639	acaaaacagtcagaagaacttggaaaccaatggaacagacttcatatctcct	2698
Db	642	#####	701
QY	2699	gtccacggaagaatctcgltgaagatgctggaataacccttggagataagctaga	2758
Db	702	#####	761
QY	2819	ctagatgatgacgcgactgtataaagctgcgacacccatgatgtctcctcattgcactc	2878
Db	762	#####	821
QY	2879	ctctgccttcccatatctgccc	2903
Db	822	#####	846
RESULT 10			
AAF58262			
ID AAF58262 standard; DNA; 936 BP.			
XX AAF58262;			
AC			
XX			
DT 24-APR-2001 (first entry)			

[illegible]

Query Match	1.9%	Score 93.4	DB 22,	Length 936;
Best Local Similarity	4.2%	Pred. No. 2.8e-16;		
Matches 34; Conservative	435;	Mismatches 336;		
		Indels 0;		Gaps 0;

Oy	2099	ttaaagtttgaagaaagaagatgactgaatcagcttcgactttgycattttaccaa	2105
	42	101
Oy	2159	gtaaaaagggaatttagctgtaactgtgagcatcttagaccaggtgaatgaactt	2218
	102	161
Oy	2219	gaatgaaatggagagcatttcgaaggacacatttgaggaaattbaacatttcta	2278
	162	221
Oy	2279	gaatccaaactggaacacaagttagctgtgtgtttcaaggccaatggaatactct	2338
	222	281
Oy	2339	agaatacctgtagcagcattgcacaactggaatccaattctagctcaattgaaatcaa	2398
	282	341
Oy	2399	aaaatgacgctctctatatactgaatccatgaatgactcgcagcatgttgagat	2458
	342	401
Oy	2459	gtcccgagttctatctatcgacagcttcaataaactatggttgaacaagttgttcac	2518
	402	461
Oy	2519	cagttgagttacaatttggagacaaagatctccttcacaggaagatggagacca	2578
		

Db 462 www..... 521
Oy 2579 aggaatccttaagttaagattccttcctccagatagaagtataaaataagagaaga 2638
Db 522 www..... 581
Oy 2639 acaaaacagtcgaagaactcttggaaccacaatggaaccagacttcatattctcct 2698
Db 582 www..... 641
Oy 2699 gtccacccgaagaatccgtgaacgaatcgtgaaattaccccttggagatcaagctaga 2758
Db 642 www..... 701
Oy 2759 gtccgagaagaagagcgaaattcttaagagagatttaattgaattggaacagcttg 2818
Db 702 www..... 761
Oy 2819 ctatagatgagcgcgaactggtataagctgcagaccatgatagtctcctcatgtgcaactc 2878
Db 762 www..... 821
Oy 2879 cctgcgccttcccatatctgcccc 2903
Db 822 cccccccccccccccccccc 846

RESULT 11

AAFS8255
ID AAF58255 standard; DNA; 938 BP.
AC AAF58255;
XX 24-APR-2001 (first entry)
XX Oligonucleotide D1876.
DE Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX Synthetic.
OS MO200107665-A2.
PN 01-FEB-2001.
PD 26-JUL-2000; 2000MO-US20476.
PE 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX (CLIN-) CLINICAL MICRO SENSORS INC.
PA Umek RM;
PI WPI: 2001-159728/16.
DR Nucleic acids containing electron-transfer group, useful as labels in
XX hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX a single surface -
PS Example 6; Page 127; 159pp; English.
XX The present invention relates to a composition comprising two nucleic
XX acids each containing an electron-transfer group (ETM) having
XX different redox potentials. The invention is used for electronic
XX detection of nucleic acids, especially of substitutions (mismatches)
XX and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
XX Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 1.9%; Score 93.4; DB 22; Length 938;
Best Local Similarity 4.2%; Pred. No. 2,8e-16;
Matches 34; Conservative 435; Mismatches 336; Indels 0; Gaps 0;

Oy 2099 ttaaagttgttagggaagaagatgacatgacatcagtcgacttgttgcatattaccaca 2158
Db 42 www..... 101
Oy 2159 gtaaaaaaaggaagtttagctgatactgttagacacatcttagacaagtgatgaagcttg 2218
Db 102 www..... 161
Oy 2219 gaatgaatggagaggtatgcaaggagccacatttgaggaaatttacaattattctta 2278
Db 162 www..... 221
Oy 2279 gaatccaaactgaaaccaaagtgtgactgtgttctcaaggccaattggaatattcct 2338
Db 222 www..... 281
Oy 2339 agaatcaactgatagcgcgaatgcacaaactggaatccagttctgacttgaattcctca 2398
Db 282 www..... 341
Oy 2399 aaaaatgagccgtccttctatccgttacctcaaccatgagtgactgtgcatgagggat 2458
Db 342 www..... 401
Oy 2459 gtcccgagttcttacttctgacagcttccaataaactatggttgaacaggtgtgtcac 2518
Db 402 www..... 461
Oy 2519 cagttgatagttacaatttttggagagcaagagatccttccaggaggaagtggagggca 2578
Db 462 www..... 521
Oy 2579 aggaatccttaagttaagattccttcctccagatagaagtataaaataagagaaga 2638
Db 522 www..... 581
Oy 2639 acaaaacagtcgaagaactcttggaaccacaatggaaccagacttcatattctcct 2698
Db 582 www..... 641
Oy 2699 gtccacccgaagaatccgtgaacgaatgctggaatcacccttgggatacagctaga 2758
Db 642 www..... 701
Oy 2759 gtccgagaagaagagcgaaattcttaagagagatttaattgaattggaacagcttg 2818
Db 702 www..... 761
Oy 2819 ctatagatgagcgcgaactggtataagctgcagaccatgatagtctcctcatgtgcaactc 2878
Db 762 www..... 821
Oy 2879 cctgcgccttcccatatctgcccc 2903
Db 822 cccccccccccccccccccc 846

RESULT 12

AAFS8252/c
ID AAF58252 standard; DNA; 936 BP.
AC AAF58252;
XX 24-APR-2001 (first entry)
XX Oligonucleotide D1835.
DE Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

THIS PAGE BLANK (USPTO)

3475 taccgggctctccagacagcccaatcaagctcccggaacagagggcgagca 3534
3300 CAGTGGCTCTCTCCAGACAGTCCGTTCAAGTACTCCGTAACGAGAGGGCCGCA 3339
3335 gctcccaagctcccaaaaggaacatttgagagaa----- 3572
3360 GCTTCTCAGCTCCACCAAGGAAAGGAACTGGAAAGATATATACAGAACATGATC 3419
3573 -----gtgctatgatatagagagagaatcgcaaatgaaact 3612
3420 CACAAGAGAGAAAGTCAAGTGGCTATGATATAGAGAGAAATCGCAAAATGAACT 3479
3613 taacaataacaaacaggtatagccgagatcagaacccagatcgagaaattacattgaa 3672
3480 TAACAATATCAAAACAGGTAGCCGGATCAGATCCAGCTGGAACATTAATTCATGCA 3539
3673 gtagctcgaatgagatggatccacatagagagagagatcgtttccataatcctcgga 3732
3540 GTATGCTCAGATGGGATTCACATAGAGGGGCAATCTGTTCTTAATATCCTCGGA 3599
3733 cagtatgtaagtatgatactcgggtttcaagagactagtagtctcgtttcagcag 3792
3600 CAGTATATGAAGGATGATCTAGCTGGTTCAAGGACTAGTAGTCTCTCGTTACAGAG 3659
3793 cacaagctacatgctccatccatcagagcgcgagagagagaaacaggaatcagttctt 3852
3660 CACAAGCTACATGCTCCGTTCAATCAGAACGCGGCAAGAGAAACAGAAATCAGTCTT 3719
3853 taccatcaaaatcaaaacagagacagatggcggtgctgggagagaaacttgacaaagac 3912
3720 TAGGTCCAAATGCAAGACGACAGATGGTGTCTCGGGGAAGACATGGCCAAAGCAC 3779
3913 cagcatcagtgagacatgctgctcactgtaagaatgacgagcagctccgacatgctc 3972
3780 CAGCATCAGTGAAGAGATGCTCAGTACGAGAGAAATGATGGCAGCTCGACACTGTC 3839
3973 agtggcgccctggtgtaacagtgagcaagagcgagatcgtatgcatgtgggcaaatgt 4032
3840 AGTGGCGCCCTTGCGACCAAGTGGCAAGAACGCGGATCTAGCATTTGGGCCAAATGGT 3899
4033 agctattgtgtgtctcagagaaagtcagtgctcctcaactcagcacaacgaaag 4092
3900 AGCTATGTGTGTTCTTCACGGAAGATGCGAGTCCCTCCAACTCAGCCAAACGGAAG 3959
4093 aggaagtaaaagctaacgagcagctgttccagaaagcagcgagacgggctcagctgga 4152
3960 AGGAGTAAAGAGTACGACACACTGCTCAGAGAGAGACGAGACCGGCTAGCACTGGA 4019
4153 gatgaagaaactgtagcccgccagcgacggcgagaaatccacatgagcagatgaaag 4212
4020 GATGAGAAACTGATGAGAACAGCAGCGGGAATCCAGGATGGCAGCAATGGAACAG 4079
4213 ctatagctcgaagaaatctgacttcccttggttcgctcgctcgtcgaacagatg 4272
4080 TTACAGCTCAAGAAAGAAATCTGATTTTCTTGGGCTCCGCTCGCTTCAGACGCAAT 4139
4273 cagtgtattctcgtatgagccttggtcctgctcagtagtggagcgcaaacctgtgtac 4332
4140 CAGCGATTTCTTGTATGATGCTGGGCTCTCAGCTAGTGGAGCGACGACCTGGCGAC 4199
4333 tctcgtcaatgggtgacattcaggtgggaatgatgataaaagggagacgttgagtaga 4392
4200 TCTTGCAATGGGTGACATTCAGTGGGGATGATGATGAAGAGGAGAGCTGAGAGTGA 4259
4393 aatcatccggcgcgcgctgtgtgttaaaacagagttccaagacactgcagcagcgt 4452
4260 GATCTTCGGGCTCAGAGCTGTGATTAACACAGGTTCCAAAGACTGCCAGACATATA 4319
4453 tctcaagtgatactcgttagacacagagctctgcatagccaaagaaagaaacagagtg 4512
4320 TGTCAAGGTGCTTACCTTTAGACAAAGAGTCTGATAGCCAAAGAAACCAAGTGGC 4379

QY 4513 gagaagaagctggagccctctgtaacagcagctctgtccttgagagagagcccccagg 4572
DB 4380 GAGAAAGAGCGTGGAGCCCCCTGTATAGCAGCTGTATCTTTTCAAGAGAGCCCCAGGG 4439
QY 4573 gaggtgttacaagatattgtctggagatattgtcgtatgataacaaatccttat 4632
DB 4440 GAGGTGTACAGATCATTTGTCTGGGAGATTATGCGCATGATCACAATCCTTTAT 4499
QY 4633 gggagtgccagatctctctatgataactggaactatccacaatgtgattgatt 4692
DB 4500 GGGAGTGGCCAGATATCTTATGATGAACTGAACTATCCAAATGTAATGATGGTT 4559
QY 4693 caaactctccctcctccctccctagatagatccaaactcgtgacactcgtgaaagagc 4752
DB 4560 CAAACTCTCCCTCTCTCCCTCTAGTAGATCCAACTTGACCTTGACAGAGAGAC 4619
QY 4753 tcccaatgctctctggaagtctaccgagacctcttaccctcgttcataagacata 4812
DB 4620 TTCCCAATGCTCTGTGAAAGTTCTACTGACCTTTTACTCTCTCATAGCACTATA 4679
QY 4813 aaactgtgtcacacaacacagcagatgcaaaacagaaagaaacagcaggtgagacc 4872
DB 4680 AACTGTGTGACAAACACAGAGCTTAAATAACAGAAAGAACACAGAGGTGG-GGAC 4738
QY 4873 cctgttaacacatgcatgctgtatgtgtgtcttaagagccagctgtaggagatacaga 4932
DB 4739 CCTGTGTCCACTGATGATGATGTTGTCTCCGAGCCCAAGTGTAGGATTAAGAGC 4798
QY 4933 agctctgtgtctcagagagag-----tctacacatgtgcccctgcaagag 4980
DB 4799 CATCTGTGTCTCTAGAGAGACCATCTGTCACACTGTGTGCTTACGCAAGG 4850

RESULT 4
AF199322 5640 bp mRNA ROD 04-JUL-2000
LOCUS AF199322 Rattus norvegicus R1M2 (R1M2) mRNA, complete cds.
DEFINITION AF199322
ACCESSION AF199322
VERSION AF199322.1 GI:8925857
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 5640)

REFERENCE 1 (bases 1 to 5640)
AUTHORS Wang, Y., Sugita, S. and Sudhof, T.C.
TITLE The R1M/R1M Family of neuronal C2 domain proteins. INTERACTIONS
WITH Rab3 AND A NEW CLASS OF SRC HOMOLOGY 3 DOMAIN PROTEINS
JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
MEDLINE 20347919
2 (bases 1 to 5640)
AUTHORS Wang, Y. and Sudhof, T.C.
REFERENCE Direct Submission
TITLE Submitted (27-OCT-1999) Center for Basic Neuroscience, The
JOURNAL University of Texas Southwestern Medical Center, 6000 Harry Hines
Blvd., Dallas, TX 75235-9111, USA

FEATURES
source
1..5640
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..5640
/gene="R1M2"
52..4719
/gene="R1M2"
/codon_start=1
/product="R1M2"
/protein_id="AA081644.1"
/db_xref="GI:8925858"
/translation="MSAPLGPGRPAAPPAASOPPEPMDLSHLYEERKIIQAVV
DRAKKEEKEESVILKHOQPEMYKQVKKMGESQOQOKGAPKPTGICHTKTFAD
QGHNCSTQOTFCARCGRSVIRSKNVKVCILCRQDILIRSGAMFVNSGNTPO
QPDKALRLRSEAPQERKAKLHEDTQFGPPGDSVPAVERKATGLTRQDSIKNG

gene
4319
CDS
4379

THIS PAGE BLANK (USPTO)

SGMKHIOIASDMPDSRKRSPSVSRPNRNYOSEEEREEXSOYVPSDSTMPRSPSOYADR
RSQREROFYEPHPLNTRDSNRGRHRSKETYIDNDEVEDSEDEYEROREEYEOYRNR
SPBNLAFYKPOPIEOMRIHAESVSRARERRHSDVSLAAVELEISLILRMDRPS
RORSVERHAAAMENORSYMERREAOQSSTPOTNHSPPRPSRIPIDRELR
ASLRKHOHLDPSSAVRKTREKMETMLRNDLSLSDPSEVRPPRHKRKKGKMR
OYLSSEEEELASTPEYSCDYETISESESRGAMGLKYVGKRPESRCLATTKVK
YVMOPEKDDRLIGTILINKRLKDGSYPRSGANMLGKLVGKRPESRCLATTKVK
KSLADTVHHRGDEVLENKRLLOGATEEVNIILESPEPOVELVSRIPIDGM
RIPDSHVALDESSSFSOSOKMOPSLISVSPMSGMRLDVPOLPLPRSRHQPSEH
TFEPVHOJIKLWFDKVGHOLIVILGKADLPSREDGPRNRYVLYFLPDSKMR
TTVKTLEPRKNOTETIYSPVHRPEPERLELEILMOARRESESELEILLELET
ALLDDEPHYKLOTHDVSSLPLPHSPYMRROLGESPPRLRKSRISSSEVSDY
CEDYGVSDVDRHARDLOSSTISVPEQVWSSNHCSPSGEHRVDVIGTRKSWSPV
PQORVBOGLRGTARGHTINTISRMDBHRYMDHYSEERDHFLLPLPRSRHQPSEH
HEDGCEADHQPHRSRSTEQPLERTTSSRSERADNTILMSPLSOSISRT
PSPALSRHPRTSVOTSPSPSVTGRGOLPOLPKGLERITEDMSTLRKNSG
ANDIEBRNOMKLNKYOVASDPRLDODYSKYSRGMDDPRGADYVSTKSDSDVSD
VSAVSTSSASRPSSTSYMSVOSERPRGNKISVTSKMSRQVGVSKSMAKSTIS
GMCSELERKDSOSDTRVAGLGTSGKRRNSIGKMAVILGSKSSASOLSQTEGG
GKLLSTVORSTETGLAVERNMRTQASRESTDGSMNSYSSEENLJFPGVRLASDQ
FSDPLDGLGPAOLVGRQTLAPAMGDIOVGMMDKQOLEVEIIRAGLIVRPGSKTLR
APYVYVLLDNGVCIAKRTKYARKLEPLEYQOLLSPEESPOGQVLOIIVMGDGRMD
HRSFPMGVAQIILDELELSNMVIGWFLFPSPSLVDPTLPLATLRASOSLESSTGSPSY
SRS"

BASE COUNT 1695 a 1301 c 1409 g 1235 t
ORIGIN

Query Match 75.8%; Score 3776; DB 94; Length 5640;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 4443; Conservative 0; Mismatches 275; Indels 422; Gaps 8;

OY 1 gcttcctagagggtgttcgctccgccaacatgtccgctccgcccggcg 60
Db 21 GCTTCCCTAGGGGTGTTCGCTCCACCAACATGTCCGCTCCACGCGCCG 80
OY 61 cccggtcccccggcggtctcctcaacctcctccgacgcccagatgtccgacgt 120
Db 81 CCGGCTCCACCGCGCGGCTCTCAGCTCTCCGACCGCGAGATGCGGACGTAG 140
OY 121 ccactcaggaagggaagaaatcctcctggtcgtcatgtgtcagaagaaga 180
Db 141 CCACCTCAGGAGAGAGAGAAATCAACGCGCTGATGATGATGATGATGATG 200
OY 181 agaaggaaggagcagtcgtgtcctcaagatcaagaagaacaagaacaacgacaca 240
Db 201 AGAGAGAGAGAGACACTCCGTCTCA----- 227
OY 241 gtggttcccttagtggatcaactgtgaataacgtctcagccacgacaaa 300
Db 228 ----- 227
OY 301 acaaccaatggaaggagccccaagaagaagctgcacacaacattgtgaatgtataaga 360
Db 228 -----AAAGCTGACATCAACAAATTGTAAGTAAGA 260
OY 361 gcaagtcagaagaatgaggaagaaatcagcagcagcagcagcagcagcagcagcagc 420
Db 261 GCAAGTCAGAGAAATGGGAGAGAAATCACACGACGACGACGAGAGAGAGAGAGAGAG 320
OY 421 gaactgtgcatctgcacacaagaacaaatttgcagatgtgtgcgcataatgttctcta 480
Db 321 GACCGCGGACATCTCCACAAAGAAATTTGCAGATGATGAGGCGCATTAACCTGTCTATA 380
OY 481 ttgcacaacccaagtctgtcgtcgtatgtgtgagtgagtgctttagctccaacaagt 540
Db 381 TTGCGCAACCAAGTTCTGT 440
OY 541 tatgtggtgtgttaattgtgtgcgaacaacaagaagaatcctcactaatacagcagatgt 600
Db 441 TATGTGGGTGTGTATTTGT 500
OY 601 gttcataatagtggttctacacactgagcaacactgataaagggttctcagagggt 660

Db 501 GTTTTACAAATAGTGCGTCTAATACACCCGACAGCCTGATCAAAAGGCTCTTCGAGGCT 560
OY 661 tcgaatgaagaaagccctcaggaagaagaacaaactacacgagcagcccgatcca 720
Db 561 TCGCAGTGAAGAACCCCTCAGGAGAAAGCAAACTGCATGAGCAACACCACTGTCCA 620
OY 721 aggaagcccggtactactacgtacccgtcaggttgagaagcgagcagcagcagcagc 780
Db 621 GGGACCCCCGGGTACCTCATAGTACCTGAGAGAGAGGTCCAGCTCATAGGCTTCA 680
OY 781 aagaacagataactataaanaatgatacagaagtgaagcaccagatgtccagtgatcgc 840
Db 681 AAGACAGGATTTCTATTAAGATGAGCTCAGAGATGAAGACACGATCCGACATGCG 740
OY 841 ttaagcagaagaacgaagtcatacgaigtgtcaggaagtcataaactgaagaatcagcaag 900
Db 741 TTGAGACGAGAAAGAAAGTCCGTAGTGTCCAGGAGATCAAAATCGAAGATGACCAAA 800
OY 901 tgaagaagaagagactacacagatgttccttcagatgtgataatgataatgataatc 960
Db 801 TGAAGAAAGAGAGATATTCACAGTATGTTCTTCAGATGACACCAATGCTTGAATCTCC 860
OY 961 ttcgattatctgatacagatctcagcgtgagccctcaatttataagaacccgtgca 1020
Db 861 ATCAGATTATGCTGATGAGAGATCTCAACGTAACCTCAATTTATGAAGACCTGATCA 920
OY 1021 ttaaatattcagagatctcacaagagagccatagacattccaaagatatgttga 1080
Db 921 TTTTAAATTATAGGATTTCTTAACAGAGAGGCGCATAGATTCACAAAGATATGTGTA 980
OY 1081 tgaatgaatgtgtgagagcagaatgtgaatataagaacaaaggagagagagagatacca 1140
Db 981 CGACGAGATGTGAGAGCGAGAGATGATGAAAGAAAGAGAGAGAGAGATGACCA 1040
OY 1141 ggcacgctacagaagtgatcacaatcctggccggtatccgttaaagccacaacccctaga 1200
Db 1041 GGCAGCCTACAGAGATGATCAAAATTTGGCCCGGTATCCGTTAAAGCCACCACTTATGA 1100
OY 1201 agaaacaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
Db 1101 AGAGCAAAATGCGGATCCACCTGAAAGTGTCCCGGCGAGACGAGAGAGAGAGAGAG 1160
OY 1261 tgttctttggcagaagcgtgaactagaagattccaggaattctcgtcgaagatgataag 1320
Db 1161 TGTTTTGGCAAAATGCTGAAGTATGCAAGATTTCTGCTTAAGATGATGATG 1220
OY 1321 accatcaagcagaagaatctgtatcgtgaagctagaagctgcaatggaagaacagagctga 1380
Db 1221 ACCATCAAGCAAGATCTGTATCTGAGCTAGAGCTGCAATGGAAGACAGCGATGCTGA 1280
OY 1381 ttcaatggaagaactcagagagcctcagggacaaagtcttataccacaagaagaccccaaa 1440
Db 1281 TTCAATGGAAGAACTCGAGAACTCAGGAGCAAAAGTTCTTATCCACAAAGAGACCAAAA 1340
OY 1441 tcatagctcccaacccctcggcgagccctataccgcttgatagaacagcagatgagcg 1500
Db 1341 TCAATCTCTCTCAACCCCTGCGAGAGCCTTATACCGCTGATGAGACCAAGCTGAGAGCG 1400
OY 1501 cgtgactccctcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560
Db 1401 TCGCGACTCCTCAAGGAAACACACCACTTATGATCCAGCTCTCTGTAGGAAAGCAAA 1460
OY 1561 gcgagaaanaatggaacacatgttaaggaatgtctttagtgcagaccagctccgagtc 1620
Db 1461 GCGAGAAANAATGGAACATGTTGAGGATGATTTTGAAGTGCAGACCGTCTGATGATC 1520
OY 1621 agtgagggccgcccacgaagcgtcataaatccaagaagaggttaaatggcgccaggt 1680
Db 1521 AGTGAAGCCGCCCCCAACAGGCTCATTAATCCAGAAAGAGAGGTAAATGCCCCAGGT 1580
OY 1681 ttcaactgagcagcctcaggaagagagctggtatcacaacccgagatatacaagctgtgata 1740

THIS PAGE BLANK (USPTO)

1531 TTCACTGAGCAGCTCAGAGAGAGAGTGGCATCCACGCTGATATACAGCTGTGATGA 1640
 1741 tgtgagcttggaaagcgaagctgtgtgtgagaaagggagacgtcaaaagggaaaaa 1800
 1641 CGTGGAGATTGAAAGCAGAGAGCTAGCTGAGAAAGAGACAT 1682
 1801 aactagtgcagggagtttgcgattcacaacaggtctgagagacaaaagaa 1860
 1683 ----- 1682
 1861 gatgtactatgtgtgcccactcttttgaagagatttgaatgtctgagccatgaat 1920
 1683 ----- GAGTACAGCTGTTGGAGCATGCTCTTG 1712
 1921 ggaacttgggtagatatacctgtagtagcacaccccttaacagagagacatgca 1980
 1713 G-----CATAGCAGTAGGCGATCCCCAATGTC 1739
 1981 taagcaccctgtgacccctgtgacgacatccaaagatggagatcgccatattgtc 2040
 1740 TTTGCAACCCGTGTGACCTGGCAGCATCCAAAGATGAGATCCCTTAATTGGCT 1799
 2041 attaaataagcgtttaaagatgtgagtgtaacctcagagactcaagagcaatgt 2100
 1800 ATTAATAAGCGTTTAAAGATGGAGTGTACCTCGAGATTCAGAGGCAATGCTGG 1859
 2101 aaaggtgtgagagaaagatgacgaatcaagctcgaacttgtgcaatttaatac 2160
 1860 AAAGGTTGTAGAGAGAAAGATCAGTCAATCAGATGCTGCAATTTATATCCAA 1919
 2161 aaaaaaggaagtttagctgtatctgtatgaacatcttaagacagggatgatac 2220
 1920 GAAAAAGGAATTTAGCTGATCTGTAGTACATCTTGAACGAGATTAACATTT 1979
 2221 atggaatgtggagcattgtgcaagggagccacatttgaagatttacaattatc 2280
 1980 ATGGAATGGAAGCGTTATTCAGAGAGCCACATTTGAGAGATTACAACTATT 2039
 2281 atcccaactgagacccaagttgagcttgttlttcaaggccaattggagatatc 2340
 2040 ATCCAAAGCCTGAAACCAAGTGTAGCTGTTGTTTCAAGCGATGAGATATGCC 2099
 2341 aatacctgtatagacgacatgacacactgtgaatcgaatttgaatttgaatc 2400
 2100 AATACCTGATAGACACTATGACACAACTGGAATCCAGTCTAGCTATTTGAAT 2159
 2401 aatggacgttcctctctatccgttactcacaatgagtcctgtgcatgtgag 2460
 2160 AATGGAATCGCCGCTATATGCTTACCTCTCCATGAGTCTCTGGCATGCTGA 2219
 2461 cccgcaattcttactctgacagcttca----- 2488
 2220 CCCACAGTCTTATCTGACACAGCTTTCAAGCCCTTATAGAAACACACGCT 2279
 2489 -----ataaactatgtgttgaacaaggttgttaccagttgattt 2532
 2280 TGTTCCTAGGGTTCAGATATAAACTATGTTTACAGGTTGCTCACCAAGTTA 2339
 2533 aatttggagcaagatctcccttccagggagaatgtggagccaaagaaatccta 2592
 2340 AATTTTGGAGCAAAAGATCTCCCTTCCAGGGAAGATGGAGGCCAAAGAAAT 2399
 2593 taagattactctctccagtagaagtgataaaataaagaagaacaaacagtc 2652
 2400 TAAAAATTATCTTCCAGACAGAGTGTATAAAACAGAGAAAGAAACAAAGTCA 2459
 2653 gaaacttggaaacccaatgtgaacagacttcaatttctctctgtccacgaaga 2712
 2460 GAAAAATTGTGGAACCAATGAAACGAGACTTTCATTTATCTCTCTGTCACG 2519
 2713 attcgttgaacgaatgtggaattaccccttggagatcaagctagagttcaga 2772
 2520 ATTCCGGGAACGAATGCTCGAAATACCTTTGGATCAAGCTCGAGTTCGAGA 2579

QY 2773 gaggcaattcttagagagagatttcaattgaatgaaacagcttctagatgagcc 2832
 Db 2580 AAGTAATTCCTTAGAGAGATTTTATGTGATTTGAAACAGCTCTGTTAGATGA 2639
 QY 2833 gcaatgtataagctgtgcaagcccaatgagtgcttctcctcattgcccctcc 2892
 Db 2640 ACACGTGTACAAATTCACAGCCCATGATGTCTCTCATTCGCTACCTCTCC 2699
 QY 2893 atactgtcccccaggagcagctccatgagagagagcccaagcagctgtcaag 2952
 Db 2700 ATATATGACGACGAGACAGCTGATGAGAGAGCCCAACAGCAGTGTGCA 2759
 QY 2953 gaaataaagtgaaagtgagtgctgtactaagctgtgaggaatgtgagtg 3012
 Db 2760 GAGAAATTAAGTACGACGAGATATCCGACTATGACGTGCGAGAGCGCGT 2819
 QY 3013 agattatcgacacaaatgtccgcgaatctcaagctccacgtgtgtgtgtgag 3072
 Db 2820 AGATTACCGACAGATGCGCGTGAATCTTCAAAAGCTCCACATTTATGAGT 2879
 QY 3073 catgtcataatcatatgtctacacacagtggtctcctcactgagtagatgta 3132
 Db 2880 CATGTCAATCAAAACATTCCTCAACATCAAGGCTCTCTCATCCAGTATGAT 2939
 QY 3133 gacaaggtcagtgctgagcctagtgtccctcctcctcaaggaatgtggaag 3192
 Db 2940 GACTAGGTATGATGTGCTAGTGTCTCTCTCTCTCTCAAGGAATGTGGAAG 2999
 QY 3193 agggacagctgtgactgtgacatlaacaaacaaatlagccgaatgtataga 3252
 Db 3000 AGGACACGCTGTACGTGGCCATTTACAAATTAACCGGATGACAGACCGT 3059
 QY 3253 ggaatgacacactctctcaatagagac----- 3280
 Db 3060 GGACGACCTTACTTCTTCAAGAGAGACAGCTATTTCTCACTTACTGCTC 3119
 QY 3281 -----aggatgtgaaagcagatagaca 3306
 Db 3120 CAGGACAGCAGTACATCCACAGAGGATGGAAGGATGTGAAGCGACGATGA 3179
 QY 3307 gccataacacagatccacatcaacagaacaaagcctctcttagagggagac 3366
 Db 3180 GCCATATCAAGATCCAGATCAACAGAAACAGGCTCTCTTACAGCGACCA 3239
 QY 3367 ctccagatcctctgacgctcgtatcaacacatcgaatgagctgagcttca 3426
 Db 3240 CTCAGATCTCTGAAAGTGTGATTAACAACTCATGAGGTGATGATTAATG 3299
 QY 3427 tggaaagatctgcccctccttcaactgtccttatacgaggtctcaccct 3486
 Db 3300 TGGAAAGATCTCCCTCTTCACTGCTTATCGAGGTCTACCCCTCGCATG 3359
 QY 3487 ccagagaaccccatcaagtlactcgygaaacaggaagggagggccagcttca 3546
 Db 3360 CCAGACAGTCCCTGAAAGTACCTCCGTATACAGAGACAGAGGCGCGAC 3419
 QY 3547 tccacaaagggagacatggagagaa----- 3572
 Db 3420 CCCACCAAGGAGAACTGTAAGATGATACAGAAAGATGAGTCCCAAGAGAG 3479
 QY 3573 -----gtgtatgagatagagagagaacacgccaatgaaacttaacaa 3624
 Db 3480 AACTCAGTGTGATGATATGAGAGAGAAATCGCAATATGAATTAACAAT 3539
 QY 3625 aaggtgagccagatcagacccacagactgtgagcaagatlaacattcgaagt 3684
 Db 3540 ACAGGTAGCCGATCAGATCCAGATCGAACAAGATTTACATCGAAGTCCG 3599
 QY 3685 atggagatccacatagagggagacatctgttccactaaatccctcgagcag 3744
 Db 3600 ATGGGATCCCATAGAGGGGCGAGTACTGTTTCTAATAATCTCTGACAGT 3659

THIS PAGE BLANK (USPTO)

[illegible]

THIS PAGE BLANK (USPTO)

File Copy
Accession No: Q9JIS1
Seq ID No: 1, Database: SPTREMBL14age 4

Db 141 QELTTSAGANFTVSGNTPOQPODKALRGLESEBAPOEKKAKLHEOTOFOGPGDSSVPA 200
QY 241 VEGRAHGLTROTITKNGSVKHOIASDMPDRKRSPSVSDONRRYQSEBREDYQYV 300
Db 201 VEGRAHGLTROTITKNGSVKHOIASDMPDRKRSPSVSDONRRYQSEBREDYQYV 260
QY 301 PSTDTMPRSPDYADRRSOREPOFEYEEPHLNYRDSNRRHSHKEYIVDEEDVESRDEY 360
Db 261 PSTDTMPRSPDYADRRSOREPOFEYEEPHLNYRDSNRRHSHKEYIVDEEDVESRDEY 320
QY 361 EROREEREYQARYRSDPNLARYPYKPOFEYEQMILHAVSARHRRSDVSLAAELED 420
Db 321 EROREEREYQARYRSDPNLARYPYKPOFEYEQMILHAVSARHRRSDVSLAAELED 380
QY 421 SRTSLRMDRPSRORASYERRAAMENORASYMERTRAOGOSYPORTSNHSPTRRSP 480
Db 381 SRTSLRMDRPSRORASYERRAAMENORASYMERTRAOGOSYPORTSNHSPTRRSP 440
QY 481 IPLDRPMRADSLARKOHNDPSSAVAKTRKREKMETMLRNDLSLSDOSEVYRPPRPHK 540
Db 441 IPLDRPMRADSLARKOHNDPSSAVAKTRKREKMETMLRNDLSLSDOSEVYRPPRPHK 500
QY 541 SKKGKMRQVSLSSSEELASTPEYSCDVELESESVSEKQKQKRTSSEGVLSDS 600
Db 501 SKKGKMRQVSLSSSEELASTPEYSCDVELESESVSEKQKQKRTSSEGVLSDS 545
QY 601 NTRSEROKRMYTGGHLEBDELESEPOIKDSGYDTCSTTLNEHSHSDKHPYWPQSK 660
Db 546 NTRSEROKRMYTGGHLEBDELESEPOIKDSGYDTCSTTLNEHSHSDKHPYWPQSK 573
QY 661 DGDRLIGRILNKLKDGSVPRDSCAMGLKLVYGGKMTESGRCAFTTKYKKSGLADTVG 720
Db 574 DGDRLIGRILNKLKDGSVPRDSCAMGLKLVYGGKMTESGRCAFTTKYKKSGLADTVG 633
QY 721 HIRPDDEVLENNGRILQAGTFEYEVNIIILSKPEPOVELYVSRPDIPIRIPDSTHAOLE 780
Db 634 HIRPDDEVLENNGRILQAGTFEYEVNIIILSKPEPOVELYVSRPDIPIRIPDSTHAOLE 693
QY 781 SSSSSFEQSKMDRPSISVSPMSPGMLRDVPOFLSGOLSTIKLMDKXGHOILYVILAKD 840
Db 694 SSSSSFEQSKMDRPSISVSPMSPGMLRDVPOFLSGOLSTIKLMDKXGHOILYVILAKD 753
QY 841 LPSRBDGRPRNRYVYKLYLPDRSDKNKRTKYKTLPEKMNQTFIYSPYHRREREML 900
Db 754 LPSRBDGRPRNRYVYKLYLPDRSDKNKRTKYKTLPEKMNQTFIYSPYHRREREML 813
QY 901 ETTIMDOARVRESESEFLEIILETALIDEPHMYKLOTHDVSSILPLRPPSPYLRQ 960
Db 814 ETTIMDOARVRESESEFLEIILETALIDEPHMYKLOTHDVSSILPLRPPSPYLRQ 873
QY 961 LHGESPTRLQSKRISDSEVSDEYDCEGVVSDYRHNCRDQSSSTLSVPEQVMSNHC 1020
Db 874 LHGESPTRLQSKRISDSEVSDEYDCEGVVSDYRHNCRDQSSSTLSVPEQVMSNHC 933
QY 1021 SPSPGPHRYDVIGTRTSMSPSAPPORANVEQHGRTATGHYNTISMDHRRVDDHYS 1080
Db 934 SPSPGPHRYDVIGTRTSMSPSAPPORANVEQHGRTATGHYNTISMDHRRVDDHYS 993
QY 1081 DRD-----RDCEADROPYHRSRSTEQRLPRLERTTSSRSER 1118
Db 994 ERDSHFLTLPRSRHROTSEHHNRDGDCEADROPYHRSRSTEQRLPRLERTTSSRSER 1053
QY 1119 PDTNLRMSPLSMGTGRSAPSPALSSHPRTGSVQTSPPSTGTGGRGROLPOLPPRGTL 1178
Db 1054 PDTNLRMSPLSMGTGRSAPSPALSSHPRTGSVQTSPPSTGTGGRGROLPOLPPRGTL 1113
QY 1179 ER-----SAMDIEBRNOMKLNKYQVAGSPRLRLEODYHSKYRSGMDHNG 1224
Db 1114 ERMTEDDOSTRKRNSGANDIEBRNOMKLNKYQVAGSPRLRLEODYHSKYRSGMDHNG 1173
QY 1225 ADYVSTKSSDSDVSVSAVSRTSSASRFSSTSYMSVOSERPRGNRKISVYTSKMNQRM 1284
Db 1174 ADYVSTKSSDSDVSVSAVSRTSSASRFSSTSYMSVOSERPRGNRKISVYTSKMNQRM 1233

QY 1285 VSGKNLTKRSTISGDKMSLEKNGSOSDTAVGALGTSCKRRRSIGAKMVAIVGSRKR 1344
Db 1234 VSGKMAKSTISGDKMSLEKNGSOSDTAVGALGTSCKRRRSIGAKMVAIVGSRKR 1293
QY 1345 SASOLSTEGGKRLRTTYVOSTETGLAVEMRNMTROASRESTGSMNSYSEGNLTFP 1404
Db 1294 SASOLSTEGGKRLRTTYVOSTETGLAVEMRNMTROASRESTGSMNSYSEGNLTFP 1353
QY 1405 GYRLASDSQSFDFLDGLGPAVLGROTLAPPMAGDIOGMMDKRGQLEVEIIRANGLYVK 1464
Db 1354 GYRLASDSQSFDFLDGLGPAVLGROTLAPPMAGDIOGMMDKRGQLEVEIIRANGLYVK 1413
QY 1465 PGSKTLPAPYVYKYLNDNGVCIARKKTKVAKTLEPLYYOQLLSEESPOGRVLOIIVMGD 1524
Db 1414 PGSKTLPAPYVYKYLNDNGVCIARKKTKVAKTLEPLYYOQLLSEESPOGRVLOIIVMGD 1473
QY 1525 YGRMDHKSFGVAQIILDELELSNMVQWFKLPFPSSLYVDTASPLTRASOSSLESSTG 1584
Db 1474 YGRMDHKSFGVAQIILDELELSNMVQWFKLPFPSSLYVDTASPLTRASOSSLESSTG 1533
QY 1585 PSYSRS 1590
Db 1534 PSYSRS 1539

RESULT 4
Q9JIS1 PRELIMINARY; PRT; 1555 AA.
AC Q9JIS1;
DT 01-OCT-2000 (TREMBLrel, 15, Created)
DT 01-OCT-2000 (TREMBLrel, 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel, 16, Last annotation update)
DE RIM2.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347919; PubMed=10748113;
RA Wang Y, Sugita S, Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins. Interactions with
Rab3 and a new class of Src homology 3 domain proteins.";
RL J. Biol. Chem. 275:20033-20044(2000).
DR EMBL: AF199322; AAF81644.1;
DR InterPro: IPR000008;
DR InterPro: IPR000306;
DR InterPro: IPR000345;
DR InterPro: IPR001478;
DR Pfam: PF00168; C2; 2;
DR Pfam: PF00595; PD2; 1;
DR PROSITE: PS50004; C2 DOMAIN 2; 2;
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1;
DR SMART: SM00239; C2; 1;
SQ SEQUENCE 1555 AA; 175912 MW; D76967BB36D9704E CRC64;

Query Match 89.6%; Score 7420.5; DB 11; Length 1555;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 1451; Conservative 21; Mismatches 31; Indels 139; Gaps 6;

QY 1 MSAPLGRGRAPAPTAASQPPPOPEMDLSHTEERKTIILAVMDRQKKEEKQSVLKI 60
Db 1 MSAPLGRGRAPAPTAASQPPPOPEMDLSHTEERKTIILAVMDRQKKEEKQSVLKI 59
QY 61 KEHKAOPTQWPFESGITELVNNVLQPOQKOPNEKEPOTKLHOOFEMYKBOVKMGESQ 120
Db 60 KEHKAOPTQWPFESGITELVNNVLQPOQKOPNEKEPOTKLHOOFEMYKBOVKMGESQ 80
QY 121 QOQEGKQAPTCGTHKTKFADGCGHNGCYCOTKFCARCGGRVSLSNKVMVQCNLCQKO 180
Db 121 QOQEGKQAPTCGTHKTKFADGCGHNGCYCOTKFCARCGGRVSLSNKVMVQCNLCQKO 171

THIS PAGE BLANK (USPTO)

```

Db 81 000EQKDAFTGICHTKTFADGCGHNCYCQYKFCACGGRHVSLSNKMVWVNLCRQ 140
Qy 181 0ELITSGAMFYNGSMTLQDPDQKVPRGLNNEAPQEKALHEDPOPOGAGDLSVA 240
Db 141 0ELITSGAMFYNGSMTLQDPDQKALGLRSEAPQEKALHEDPOGAGDLSVA 200
Qy 241 0VEGRAHGLTFKODITKNGSGVGHQIASDMPDRKSPSVSDQNRREYQSEEREDSQY 300
Db 201 0VEGRAHGLTFKODITKNGSGVGHQIASDMPDRKSPSVSDQNRREYQSEEREDSQY 260
Qy 301 0PSDGTMRSPSDYADRSGQREPOFEERPHLNYRDSNRGRHRSKEYIVDEDEVESRDEY 360
Db 261 0PSDGTMRSPSDYADRSGQREPOFEERPHLNYRDSNRGRHRSKEYIVDEDEVESRDEY 320
Qy 361 0RQREBEYQARYSDPLARYPVKQPYEEDMRTHAEYSRAREHRSDDVSLAMELE 420
Db 321 0RQREBEYQARYSDPLARYPVKQPYEEDMRTHAEYSRAREHRSDDVSLAMELE 380
Qy 421 0SRILHMDRPSRQSVERRAAMNORSYSMEPTREAGOGSSYPORTNHSPPTRSP 480
Db 381 0SRILHMDRPSRQSVERRAAMNORSYSMEPTREAGOGSSYPORTNHSPPTRSP 440
Qy 481 0PLDRPMRRADSLRKQHHLDPSAVERKTKREKMETMLRNDLSLSDQSESVRPPRPHK 540
Db 441 0PLDRPMRRADSLRKQHHLDPSAVERKTKREKMETMLRNDLSLSDQSESVRPPRPHK 500
Qy 541 0SKGCKARQVSLSESEELASTPEYTCSDVLESESVSEKQSGKRTSDQGLSDS 600
Db 501 0SKGCKARQVSLSESEELASTPEYTCSDVLESESVSEKQSGKRTSDQGLSDS 560
Qy 601 0NTRBERQKRRNYGSHSLDEEDLEWSEPOIKDQGVPTCSSTTLNEHSHSDKHPTWQPSK 660
Db 561 0NTRBERQKRRNYGSHSLDEEDLEWSEPOIKDQGVPTCSSTTLNEHSHSDKHPTWQPSK 520
Qy 661 0DGRILGRILNKLKDKSVPRDSCAMGLKLVGCKMTESGRLCAFTTKYKKSGLADTVG 720
Db 574 0DGRILGRILNKLKDKSVPRDSCAMGLKLVGCKMTESGRLCAFTTKYKKSGLADTVG 633
Qy 721 0HLRDEVLWENGLQATFEVYNIILLESKPEPEYELVNSRPIDIPRIPOSTHAOLE 780
Db 634 0HLRDEVLWENGLQATFEVYNIILLESKPEPEYELVNSRPIDIPRIPOSTHAOLE 693
Qy 781 0SSSSSEFESQKDRPSISVTPSPSPGMLRDVPOFLSGOLS-----IKLME 824
Db 694 0SSSSSEFESQKDRPSISVTPSPSPGMLRDVPOFLSGOLS-----IKLME 753
Qy 825 0KVGHQLIVTILGANDLPREDGRPNRYKTYFLDRSDKNRRTKTKLEPKWNO 884
Db 754 0KVGHQLIVTILGANDLPREDGRPNRYKTYFLDRSDKNRRTKTKLEPKWNO 813
Qy 885 0FYSVHRHREPERKMLLETLMDQARVRESESEFLGELLLELALDDEPHWYKLOTHDV 944
Db 814 0FYSVHRHREPERKMLLETLMDQARVRESESEFLGELLLELALDDEPHWYKLOTHDV 873
Qy 945 0SLPRLPRSPYPLRRLQHGESPTRLQORSKRISDSEVSDYDCEDGVVNSDYRHHNGDLQ 1004
Db 874 0SLPRLPRSPYPLRRLQHGESPTRLQORSKRISDSEVSDYDCEDGVVNSDYRHHNGDLQ 933
Qy 1005 0SSTLSVPEQVMSNHCSPSGSPHARDVIGRTSRWSBSPAPPRQNRVEOGRHGTATGHNT 1064
Db 934 0SSTLSVPEQVMSNHCSPSGSPHARDVIGRTSRWSBSPAPPRQNRVEOGRHGTATGHNT 993
Qy 1065 0ISMRDRHVMDDHYSDDP-----RCEADADROPYHRSRSTEQ 1102
Db 994 0ISMRDRHVMDDHYSDDP-----RCEADADROPYHRSRSTEQ 1053
Qy 1103 0PLERTRTSSSESPEDNLMRSMPLMTGRSAPSPALSRSHPTGVSQVOTSSSTPGT 1162
Db 1034 0PLERTRTSSSESPEDNLMRSMPLMTGRSAPSPALSRSHPTGVSQVOTSSSTPGT 1113
Qy 1163 0GRRGQLPOLPKGTLE-----SAMDIEERNRQMLNKYKOVAGSDPRL 1208
Db 1114 0GRRGQLPOLPKGTLE-----SAMDIEERNRQMLNKYKOVAGSDPRL 1173

```

```

Qy 1209 0QYHSHKYSRSGMDPRGADTVSTKSSDSVDVSAVSRTSSASRFSSTSYMVSQSERPRGN 1268
Db 1174 0QYHSHKYSRSGMDPRGADTVSTKSSDSVDVSAVSRTSSASRFSSTSYMVSQSERPRGN 1233
Qy 1269 0KISVFTSKMONROMGVSGKMLTKSTISISGDMCSLEKNDGSDOPRAGALGTSGKRPRS 1328
Db 1234 0KISVFTSKMONROMGVSGKMLTKSTISISGDMCSLEKNDGSDOPRAGALGTSGKRPRS 1293
Qy 1329 0IGAKVAVIYGLSRKSRSSASOLSOTEGGKKLRSYVORSTETGLAVEMRMNTQOASREST 1388
Db 1294 0IGAKVAVIYGLSRKSRSSASOLSOTEGGKKLRSYVORSTETGLAVEMRMNTQOASREST 1353
Qy 1389 0DGSNMSYSSSEGNLIFPGYRLASDSQFSDFLDGLGPAOLVGRQTLATPANGDITGVMMDK 1448
Db 1354 0DGSNMSYSSSEGNLIFPGYRLASDSQFSDFLDGLGPAOLVGRQTLATPANGDITGVMMDK 1413
Qy 1449 0GLEVEITIRANGLVVPGSKTLPADYKYYLLDNGVCIAKKTKVARKTLEPIYQQLLSF 1508
Db 1414 0GLEVEITIRANGLVVPGSKTLPADYKYYLLDNGVCIAKKTKVARKTLEPIYQQLLSF 1473
Qy 1509 0ESPOGRVLOIIVWGDYGRMDKSPMGVAQIILDELSNMVYGMFKLPPSSLYDPTSA 1568
Db 1474 0ESPOGRVLOIIVWGDYGRMDKSPMGVAQIILDELSNMVYGMFKLPPSSLYDPTSA 1533
Qy 1569 0PLTRASQSSLESSTGSPYSRS 1590
Db 1534 0PLTRASQSSLESSTGSPYSRS 1555

```

```

RESULT 5
09JXSO PRELIMINARY: PRT: 1399 AA.
ID 09JXSO
AC 09JXSO
DT 01-OCT-2000 (TREMBLrel.145, Created)
DT 01-OCT-2000 (TREMBLrel.15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel.16, Last annotation update)
DE RIM2-XB.
GN RIM2-2B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu.; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE-203479.9; PubMed-10748113;
RA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins: Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins."
RL J. Biol. Chem. 275:20033-20044(2000).
DR EMBL: AF199324; AAF01646.1;
DR InterPro: IPR000008;
DR InterPro: IPR000306;
DR InterPro: IPR000345;
DR InterPro: IPR001478;
DR Pfam: PF00168; C2; 2.
DR Pfam: PF00595; PDZ; 1.
DR PROSITE: PS00004; C2 DOMAIN; 2.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART: SM00239; C2; 1.
SQ SEQUENCE 1399 AA; 158945 MW; 3F1550E3D52D237B CRC64;

```

```

Query Match 82.8%; Score 6856.5; DB 11; Length 1399;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 138; Conservative 17; Mismatches 22; Indels 235; Gaps 3;
Qy 1 MSAPLGRGRPATPPAASQPPPEMPDLSHLTEEEKIILAVNDROKKEEKSQVYKI 60
Db 1 MSAPLGRGRPATPPAASQPPPEMPDLSHLTEEEKIILAVNDROKKEEKSQVYK 59
Qy 61 KEEHKAQPTQWPEFSSITELVNNVLOPQOKQPNKEKEPQTKLHOFEMAYKEVYKQGESQ 120
Db 61 KEEHKAQPTQWPEFSSITELVNNVLOPQOKQPNKEKEPQTKLHOFEMAYKEVYKQGESQ 119

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:05:34 ; Search time 43.1 seconds
(without alignments)
4880.858 Million cell updates/sec

Title: US-09-617-099b-1
Perfect score: 8285
Sequence: 1 MSAPLPGRPAPTPAASQP.....TRRASQSSLESSTGPGYSRS 1590

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 13205027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_protent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7930	95.7	1530	11 09EQZ7	09eqz7 mus muscula
2	7459.5	90.0	1525	11 09JIR5	09jir5 rattus norv
3	7438.5	89.8	1539	11 09JIR6	09jir6 rattus norv
4	7420.5	89.6	1553	11 09JIR6	09jir6 rattus norv
5	6856.5	82.8	1399	11 09JIR5	09jir5 rattus norv
6	6543	79.0	1330	11 09JIR7	09jir7 rattus norv
7	6522	78.7	1352	11 09JIR7	09jir7 rattus norv
8	6504	78.5	1368	11 09JIR9	09jir9 rattus norv
9	6212	75.0	1292	11 09JIR8	09jir8 rattus norv
10	5638	68.1	1188	4 09UQ26	09uq26 homo sapien
11	4513.5	54.5	1615	11 09JIR4	09jir4 rattus norv
12	4489.5	54.2	1553	11 035168	035168 rattus norv
13	2936.5	35.4	1053	4 015048	015048 homo sapien
14	2424.5	29.3	740	4 09HBA5	09hba5 homo sapien
15	2410.5	28.1	766	4 09HBA6	09hba6 homo sapien
16	2340.5	28.2	700	4 09HBA4	09hba4 homo sapien
17	1912	23.1	369	4 043413	043413 homo sapien
18	1834.5	22.1	596	4 09HBA3	09hba3 homo sapien
19	1821	22.0	567	4 09HBA2	09hba2 homo sapien

20	1763.5	21.3	516	4 09HBA1	09hba1 homo sapien
21	1507.5	18.2	471	4 09HBA3	09hba3 homo sapien
22	1221.5	14.7	285	11 09JIR2	09jir2 rattus norv
23	1025.5	12.4	308	4 09JUD0	09jud0 homo sapien
24	1022	12.3	307	11 09JIR3	09jir3 rattus norv
25	1021.5	12.3	308	4 092511	092511 homo sapien
26	1002	12.1	853	5 022366	022366 caenorhabdi
27	989	11.9	2464	5 09VEE5	09vee5 drosophila
28	842	10.2	256	4 09H426	09h426 homo sapien
29	637	7.7	5120	13 09PU36	09pu36 gallus gall
30	555	6.7	5085	11 09JKS6	09jks6 rattus norv
31	546.5	6.6	5038	11 09QYX7	09qyx7 mus muscula
32	500.5	6.0	743	4 043373	043373 homo sapien
33	475.5	5.7	1212	4 060305	060305 homo sapien
34	469	5.7	4880	11 09JIR1	09jir1 rattus norv
35	459.5	5.5	4833	11 09QYX6	09qyx6 mus muscula
36	437	5.3	1325	5 094227	094227 caenorhabdi
37	371	4.5	2296	4 09UHA8	09uha8 homo sapien
38	369	4.3	2752	4 09UQ35	09uq35 homo sapien
39	358	4.3	547	5 021163	021163 caenorhabdi
40	324.5	3.9	1262	4 09UQ40	09uq40 homo sapien
41	295	3.6	3111	5 09VH10	09vh10 drosophila
42	291.5	3.5	2951	5 09W320	09w320 drosophila
43	290.5	3.5	5476	5 09NJ17	09nj17 drosophila
44	290.5	3.5	5533	5 09U6C3	09u6c3 drosophila
45	290.5	3.5	5533	5 09VPL2	09vpl2 drosophila

ALIGNMENTS

RESULT 1

ID	09EQZ7	PRELIMINARY:	PRT: 1530 AA.
AC	09EQZ7;		
DT	01-MAR-2001 (TREMBLrel. 16, created)		
DT	01-MAR-2001 (TREMBLrel. 16, last sequence update)		
DT	01-MAR-2001 (TREMBLrel. 16, last annotation update)		
DE	RIM2.		
GN	RIM2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	pubMed-11056535;		
RA	Ozaki N., Shibasaki T., Kashima Y., Miki T., Takahashi K., Ueno H., Sunaga Y., Yano H., Matsura Y., Iwanaga T., Takai Y., Saito S.;		
RT	"cAMP-GEFII is a direct target of cAMP in regulated exocytosis."		
RL	Nat. Cell Biol. 2:805-811(2000).		
DR	EMBL; AB021131; BAB18975.1;		
SO	SEQUENCE 1530 AA; 172863 MW; 58CF11BF7152357D CRC64;		

Query Match	95.7%	Score 7930;	DB 11;	Length 1530;
Best Local Similarity	96.2%	Pred. No. 0;		
Matches 1529;	Conservative	0;	Mismatches 1;	Indels 60;
				Gaps 1;
QY	1	MSAPLPGRPAPTPAASQPPPEMDLSHLTEEEKITLAWDRKKEEKQSYLKI	60	
DB	1	MSAPLPGRPAPTPAASQPPPEMDLSHLTEEEKITLAWDRKKEEKQSYLKI	60	
QY	61	KEEHKAQPTQWPFSGITELVNNVLOPQOKOPNEKEBPQTLHOOFEMYKQVKKMGESQ	120	
DB	61	KEEHKAQPTQWPFSGITELVNNVLOPQOKOPNEKEBPQTLHOOFEMYKQVKKMGESQ	120	
QY	121	OOOEKGAAPTCGICHTKTEFAGCGHNCSCQKRCFACGGRVSLRNNKMMWCNLCRKQ	180	
DB	121	OOOEKGAAPTCGICHTKTEFAGCGHNCSCQKRCFACGGRVSLRNNKMMWCNLCRKQ	180	
QY	181	OEILTKSGAMFYNSGNTLQOPDQKVPRLRNEAPQEKRAKLHEOPQFGAPDLSVPA	240	

```

Db 181 QEILTKSGAMFYNSGSENTLOQPDQKVPGRGLRNEAPQOEKKAKLHEQPOFGAGDLSTVPA 240
QY 241 VEGRAHGLTRQDITKNGSVKHOIASDMPSPKRKSPSVSRQONRREYEDSEEREDYSQYV 300
Db 241 VEGRAHGLTRQDITKNGSVKHOIASDMPSPKRKSPSVSRQONRREYEDSEEREDYSQYV 300
QY 301 PSDGTPMRSPSDYADRRSQREPOFEYEEPGHLNYPDSNRGRHHSKEYIYDDDEDVESRDEY 360
Db 301 PSDGTPMRSPSDYADRRSQREPOFEYEEPGHLNYPDSNRGRHHSKEYIYDDDEDVESRDEY 360
QY 361 ERQREEREYQARYSDNRLAYPYKPPQYEEOMRIHAEVSRAHERRHSDVSLANAELDD 420
Db 361 ERQREEREYQARYSDNRLAYPYKPPQYEEOMRIHAEVSRAHERRHSDVSLANAELDD 420
QY 421 SRSLILMDPPRSRQSVSERRAAMENORSYMERTRREAOGOSYTORPNSHSPPTPRSP 480
Db 421 SRSLILMDPPRSRQSVSERRAAMENORSYMERTRREAOGOSYTORPNSHSPPTPRSP 480
QY 481 IPLDRPDMRRADSLRKQHLLDPSSAVRKTREKMETMLRNDLSLSDQSESVAPPPRPBK 540
Db 481 IPLDRPDMRRADSLRKQHLLDPSSAVRKTREKMETMLRNDLSLSDQSESVAPPPRPBK 540
QY 541 SKKGKMKROYSLSSSEELASTPEYTCDDVLELSSEYSEKGDQSKRKRTSEQGVLSDS 600
Db 541 SKKGKMKROYSLSSSEELASTPEYTCDDVLELSSEYSEKGDQSKRKRTSEQGVLSDS 600
QY 601 NTRSERQKKEMYGGHSLLEDLEWSEPOIKDSGYDTCSSTLNEESHSDKHPPVWQPSK 660
Db 601 NTRSERQKKEMYGGHSLLEDLEWSEPOIKDSGYDTCSSTLNEESHSDKHPPVWQPSK 660
QY 661 DGRLLIGRILLNKLKDGSPVRODSGAMGLKVGKMTESGRLCAFITKVKKGLADTVG 720
Db 661 DGRLLIGRILLNKLKDGSPVRODSGAMGLKVGKMTESGRLCAFITKVKKGLADTVG 720
QY 721 HLRPGGEVLEWNRGLLOGAFEEYNNIILLESKPEPOVELVYSPRGIDPRIPDSTHAOLE 780
Db 721 HLRPGGEVLEWNRGLLOGAFEEYNNIILLESKPEPOVELVYSPRGIDPRIPDSTHAOLE 780
QY 781 SSSSESESOAMDPRPISVTPSPMSPGMLRDVPOFLSGOLSIKMFDPKVGHQILVITLGAKD 840
Db 781 SSSSESESOAMDPRPISVTPSPMSPGMLRDVPOFLSGOLSIKMFDPKVGHQILVITLGAKD 840
QY 841 LPRREGGRPNRPVYKTYFLDRDSKSNKRRTKYVAKTLEPKKNQTFIYSPVHRREREML 900
Db 841 LPRREGGRPNRPVYKTYFLDRDSKSNKRRTKYVAKTLEPKKNQTFIYSPVHRREREML 900
QY 901 ETLIMDOARVRESESEFLGILIELEFALLDDPEHMYKLOTHDVSSPLPRPSPYLPBRQ 960
Db 901 ETLIMDOARVRESESEFLGILIELEFALLDDPEHMYKLOTHDVSSPLPRPSPYLPBRQ 960
QY 961 LHGESPTRLQRSKRISDSSEVSDYDCEDGVGVVSDYRHNGRDLOSTLSVPEQVWSSNHC 1020
Db 961 LHGESPTRLQRSKRISDSSEVSDYDCEDGVGVVSDYRHNGRDLOSTLSVPEQVWSSNHC 1020
QY 1021 SPGSGPHRYDVIGRTKSWSPAPPPQNNVQOGRGTATGHYNTISAMDHRVVDHHS 1080
Db 1021 SPGSGPHRYDVIGRTKSWSPAPPPQNNVQOGRGTATGHYNTISAMDHRVVDHHS 1080
QY 1081 DHRDRCDEADROPYHRSRSTQRPLEERTTTRSSSERPDTNLMRSMPLMTGSAVPS 1140
Db 1081 DHRDRCDEADROPYHRSRSTQRPLEERTTTRSSSERPDTNLMRSMPLMTGSAVPS 1140
QY 1141 ALSRSHPTGVSOTSPSSPTPGTGRGRQOLPQLPPKGTLEERSAMDIEERNQMLNKYKOV 1200
Db 1141 ALSRSHPTGVSOTSPSSPTPGTGRGRQOLPQLPPKGTLEERSAMDIEERNQMLNKYKOV 1200
QY 1201 AGSDRLLEDYSHKYSKGMDDPHRGADTVSTKSSDSYDVSAVSRSTSSASRSTSTMSV 1260
Db 1201 AGSDRLLEDYSHKYSKGMDDPHRGADTVSTKSSDSYDVSAVSRSTSSASRSTSTMSV 1260
QY 1261 QSERPRGNRKISVFTSKMONRQMGVSGKNLTKSTISGDMCSLEKNDGSDSDTAVGLGT 1320
Db 1261 QSERPRGNRKISVFTSKMONRQMGVSGKNLTKSTISGDMCSLEKNDGSDSDTAVGLGT 1320
QY 1321 QSERPRGNRKISVFTSKMONRQMGVSGKNLTKSTISGDMCSLEKNDGSDSDTAVGLGT 1380
Db 1321 QSERPRGNRKISVFTSKMONRQMGVSGKNLTKSTISGDMCSLEKNDGSDSDTAVGLGT 1380
QY 1381 ROASRESTDGSMSYSEGNLIFPGVRLASQSFDFLDGCPAQOLVGRQTLAPAMDI 1440
Db 1381 ROASRESTDGSMSYSEGNLIFPGVRLASQSFDFLDGCPAQOLVGRQTLAPAMDI 1440
QY 1441 QVGMDDKKGQLEVEILIRAGLVYKPGSKTLPAPYVKYLLDNGVCIKKTKRVARTLEP 1500
Db 1441 QVGMDDKKGQLEVEILIRAGLVYKPGSKTLPAPYVKYLLDNGVCIKKTKRVARTLEP 1500
QY 1501 LYOOQLSFEESSPGGRVLIQIIVWGDYGRMDKSPFNGVQOILDELIELSNMVTGFKLPSPS 1560
Db 1501 LYOOQLSFEESSPGGRVLIQIIVWGDYGRMDKSPFNGVQOILDELIELSNMVTGFKLPSPS 1560
QY 1561 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1590
Db 1501 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1590
QY 1590 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1650
Db 1501 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1650

```

```

QY 1321 SGKKRRSSIGAKKVAIVGLSKRSASQASQOTEGGKKLRSTVQSTETGLAVENRMNT 1380
Db 1261 SGKKRRSSIGAKKVAIVGLSKRSASQASQOTEGGKKLRSTVQSTETGLAVENRMNT 1320
QY 1381 ROASRESTDGSMSYSEGNLIFPGVRLASQSFDFLDGCPAQOLVGRQTLAPAMDI 1440
Db 1321 ROASRESTDGSMSYSEGNLIFPGVRLASQSFDFLDGCPAQOLVGRQTLAPAMDI 1380
QY 1441 QVGMDDKKGQLEVEILIRAGLVYKPGSKTLPAPYVKYLLDNGVCIKKTKRVARTLEP 1500
Db 1381 QVGMDDKKGQLEVEILIRAGLVYKPGSKTLPAPYVKYLLDNGVCIKKTKRVARTLEP 1440
QY 1501 LYOOQLSFEESSPGGRVLIQIIVWGDYGRMDKSPFNGVQOILDELIELSNMVTGFKLPSPS 1560
Db 1441 LYOOQLSFEESSPGGRVLIQIIVWGDYGRMDKSPFNGVQOILDELIELSNMVTGFKLPSPS 1500
QY 1561 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1590
Db 1501 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1590

RESULT 2
Q9JIR5 PRELIMINARY; PRT; 1525 AA.
AC Q9JIR5:
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, last annotation update)
DE RIM2-5B.
GN RIM2-5B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347919; PubMed=10748113;
RA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins. Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins.";
RL J. Biol. Chem. 275:20033-20044(2000).
DR EMBL; AF199331; AAF81653.1; -.
DR InterPro; IPR000008; -.
DR InterPro; IPR000306; -.
DR InterPro; IPR000345; -.
DR InterPro; IPR001478; -.
DR Pfam; PF00168; C2; 2.
DR Pfam; PF00595; PD2; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 2.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART; SM00239; C2; 1.
SQ SEQUENCE 1525 AA; 172435 MW; D4930D62F03CF21 CRC64;

Query Match 90.0%; Score 7459.5; DB 11; Length 1525;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 1452; Conservative 21; Mismatches 30; Indels 109; Gaps 4;

QY 1 MSAPLGRGPAPTPAASQPPQPEMDLSHLTEERKILAYMDROKKEEKEQSVLKI 60
Db 1 MSAPLGRGPAPTPAASQPPQPEMDLSHLTEERKILAYMDROKKEEKEQSVLKI 59
QY 61 KEENKAQPTQWPFPGSITLVANNVLOPQOKOPREKEPQTKLHQEFMYKEQYKMEESQ 120
Db 60 -----KLHQEFMYKEQYKMEESQ 80
QY 121 QOEOGKDAPTGICHTKTFADGCHNCSCYCTKFCARCGGRVLSNKNVYWCNLCRRO 180
Db 81 QOEOGKDAPTGICHTKTFADGCHNCSCYCTKFCARCGGRVLSNKNVYWCNLCRRO 140
QY 181 QEILTKSGAMFYNSGSENTLOQPDQKVPGRGLRNEAPQOEKKAKLHEQPOFGAGDLSTVPA 240
Db 181 QEILTKSGAMFYNSGSENTLOQPDQKVPGRGLRNEAPQOEKKAKLHEQPOFGAGDLSTVPA 240

```

```
Db 141 QELLTSGAMFYNSGNTPEQPODOKALRGLRSEAPQOEKAKLHEOTQOFQPGSDSVPA 200
Qy 241 VEKGAHGLTROTDTIKNSGVKHOIASDMPDCKRSPSVSRDONRRYEOSEEDYQOYV 300
Db 201 VERGAHGLTROTDTIKNSGVKHOIASDMPDCKRSPSVSRDONRRYEOSEEDYQOYV 260
Qy 301 PSDGTMPSPSDYADRRSOREPOFEPECHLNRDSNRGRHRSKEYITVDEDEVSDEY 360
Db 261 PSDGTMPSPSDYADRRSOREPOFEPECHLNRDSNRGRHRSKEYITVDEDEVSDEY 320
Qy 361 ERORREERYQARYRSDPNLARYPVKQPYEEOMRIHAENSRARHRRHSDVSLANAELED 420
Db 321 ERORREERYQARYRSDPNLARYPVKQPYEEOMRIHAENSRARHRRHSDVSLANAELED 380
Qy 421 SRISLLRMDRPRORSVSRRAAMENORSTSMERTREAOSSYPORTSNHSPPTPRSP 480
Db 381 SRISLLRMDRPRORSVSRRAAMENORSTSMERTREAOSSYPORTSNHSPPTPRSP 440
Qy 481 IPLDRPMRADSLRKOHILDPSSAVAKTKREKMETMLRNDLSLSDQSESVRPPRPBK 540
Db 441 IPLDRPMRADSLRKOHILDPSSAVAKTKREKMETMLRNDLSLSDQSESVRPPRPBK 500
Qy 541 SKKGKMRQVSLSSSEBELASTPEYTSQDVELESSESVSEKQSGKRRKTSSEQVLSDS 600
Db 501 SKKGKMRQVSLSSSEBELASTPEYTSQDVELESSESVSEKQSGKRRKTSSEQVLSDS 560
Qy 601 NTRSRQKRMVYGGHSLDELEDLEWSEPOIKDQGVDRCSSTLMEHSHSDKHPTWQPSK 660
Db 566 -----YSLWEHASWH-----SSEASPMSLHPPTWQPSK 573
Qy 661 DGDRLIGRILLKRLKDGSVPRDSCGMLGLKVGCKMTESGRLCAFTYKKKSLADTVG 720
Db 574 DGDRLIGRILLKRLKDGSVPRDSCGMLGLKVGCKMTESGRLCAFTYKKKSLADTVG 633
Qy 721 HLRPQDEVLEWNGRLQGAFTFEEVYNIIESEKPEVOYELVSRPQDIPRIPISTHAOLE 780
Db 634 HLRPQDEVLEWNGRLQGAFTFEEVYNIIESEKPEVOYELVSRPQDIPRIPISTHAOLE 693
Qy 781 SSSSFESEQKMRPSTSVSPMSGMLRQVPOFLSGOLSTIKLWFDKQGHOLITVILAKD 840
Db 694 SSSSFESEQKMRPSTSVSPMSGMLRQVPOFLSGOLSTIKLWFDKQGHOLITVILAKD 753
Qy 841 LPSRDGGRPRNRYVYKIFLPRDSDDKRRKRTKYTKLEPMNQTFFISPVYHRREREML 900
Db 754 LPSRDGGRPRNRYVYKIFLPRDSDDKRRKRTKYTKLEPMNQTFFISPVYHRREREML 813
Qy 901 ETLMDQARVRESEEFELGELLIELETALLDDEPHWKIQTHDVSSLLPRPSPYLPRRO 960
Db 814 ETLMDQARVRESEEFELGELLIELETALLDDEPHWKIQTHDVSSLLPRPSPYLPRRO 873
Qy 961 LHGESPTRLQSKRISDSDEVSVDYDCEDEGVGVSDYRHNGRDLOSSLVSLVPEQVMSNHC 1020
Db 874 LHGESPTRLQSKRISDSDEVSVDYDCEDEGVGVSDYRHNGRDLOSSLVSLVPEQVMSNHC 933
Qy 1021 SPSSGPHRVDVIGRTSRSPSPAPPQORANYEOGHRGTATGHYNTISPMDRHRYVMDHXS 1080
Db 934 SPSSGPHRVDVIGRTSRSPSPAPPQORANYEOGHRGTATGHYNTISPMDRHRYVMDHXS 993
Qy 1081 DRD-----RDCEAADRQYHRSRSTEDRPLLERTTTRSSER 1118
Db 994 ERDSSHFLTLPRSRHROTSEHHNHRDGCDEAADRQYHRSRSTEDRPLLERTTTRSSER 1053
Qy 1119 PDTNLMRSPSLMTGTSAPSPALSRSHPTGTSVOTSPSSPTPGTGRGRQOLPOLPRGTL 1178
Db 1054 ADTNLMRSPSLMTGTSAPSPALSRSHPTGTSVOTSPSSPTPGTGRGRQOLPOLPRGTL 1113
Qy 1179 ERSSADIERNOMKLINKKOYVAGSDPRLQDYHSHKYRSGMDPHRQADTVSTKSSDSVYS 1238
Db 1114 ERSSADIERNOMKLINKKOYVAGSDPRLQDYHSHKYRSGMDPHRQADTVSTKSSDSVYS 1173
Qy 1239 DVSASVRSSTSSASRFSSTYSVOSVSEPRRGNRKISVFTSKMNQMGVSGKNLTKSTISG 1298
Db 1174 DVSASVRSSTSSASRFSSTYSVOSVSEPRRGNRKISVFTSKMNQMGVSGKNLTKSTISG 1233
```

```
Qy 1299 DMCISLEKNDGSDPFAVAGALGTSGKKRRSSIGAKMAIVGLSRKSRASQLSQTEGGCK 1358
Db 1234 DMCISLEKNDGSDPFAVAGALGTSGKKRRSSIGAKMAIVGLSRKSRASQLSQTEGGCK 1293
Qy 1359 LRSTVORSTETGLAVEMRMWMTROASRESTDGSMSNSSEGNLFFPGVRLASDSQSDFL 1418
Db 1294 LRSTVORSTETGLAVEMRMWMTROASRESTDGSMSNSSEGNLFFPGVRLASDSQSDFL 1353
Qy 1419 DGLGPAQLVGRQTLATPAGDIQVGMADKQGLEVEILIRAGLVYKSGSKLPPAPYKYV 1478
Db 1354 DGLGPAQLVGRQTLATPAGDIQVGMADKQGLEVEILIRAGLVYKSGSKLPPAPYKYV 1413
Qy 1479 LLDNGVCIAKKKTVAARKTLEPLYOOLLSPEESPQGRVLOIIVYGDYGRMDHKSGFMGYAO 1538
Db 1414 LLDNGVCIAKKKTVAARKTLEPLYOOLLSPEESPQGRVLOIIVYGDYGRMDHKSGFMGYAO 1473
Qy 1539 ILDELELSNMVIGWFKLPPSSLYVDPTSAPLTRRASQSSLESSTGSPSRS 1590
Db 1474 ILDELELSNMVIGWFKLPPSSLYVDPTSAPLTRRASQSSLESSTGSPSRS 1525

RESULT 3
QyIR6 PRELIMINARY; PRT; 1539 AA.
AC QyIR6;
DT 01-OCT-2000 (TREMBLREL, 15, Created)
DT 01-OCT-2000 (TREMBLREL, 15, Last sequence update)
DT 01-MAR-2001 (TREMBLREL, 16, Last annotation update)
DE RIM-5A.
GN RIM-5A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347919; PubMed=10748113;
RA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins. Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins.";
RL J. Biol. Chem. 275:20033-20044(2000).
DR EMBL: AF199330; AAF81652.1; -.
DR InterPro: IPR000008; -.
DR InterPro: IPR000306; -.
DR InterPro: IPR000345; -.
DR InterPro: IPR001478; -.
DR Pfam: PF00168; C2; 2.
DR Pfam: PF00595; PDZ; 1.
DR PROSITE: PS00004; C2_DOMAIN_2; 2.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART: SM00239; C2; 1.
SQ SEQUENCE 1539 AA; 174071 MW; 12882D2F4E0AE8C CRC64;

Query Match 89.8%; Score 7438.5; DB 11; Length 1539;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1451; Conservative 21; Mismatches 31; Indels 123; Gaps 5;
```

```
Db 141 OEILTFSGAMFYNSGNTPOQDQKALRGLRSEAEAPQEKAKLHEQTOFOQGPBDSVPA 200
QY 241 VEGRAHAGLIRQOTIKNGSGVKHOLASDMPDRKRSVSDQNNRYEQSEREDYSQYV 300
Db 201 VEGRAHAGLIRQOTIKNGSGVKHOLASDMPDRKRSVSDQNNRYEQSEREDYSQYV 260
QY 301 PSDGTPRSPSDVADRKSOREPOFYEEPGHLNRYDSNRGRHRSKEYIYDDDDVSRDEY 360
Db 261 PDSOTMPRSPSDVADRKSOREPOFYEEPDHLNRYDSNRGRHRSKEYIYDDDDVSRDEY 320
QY 361 ERORREBEYOARYRSDENLARYPVKPOPYEEOMRIHAEVSAARHERHRSDVSLANAELTD 420
Db 321 ERORREBEYOARYRSDENLARYPVKPOPYEEOMRIHAEVSAARHERHRSDVSLANAELTD 380
QY 421 SRSLIRMDPSRORSYSEARAAMENORSYMERTRERAOGOSYQKRNHSPPTPRSP 480
Db 381 SRSLIRMDPSRORSYSEARAAMENORSYMERTRERAOGOSYQKRNHSPPTPRSP 440
QY 481 IPLDRPDMRADSLRKHHLDPSSAVRKTREKMETMLRNDLSLSDGSESVPRPPRPK 540
Db 441 IPLDRPDMRADSLRKHHLDPSSAVRKTREKMETMLRNDLSLSDGSESVPRPPRPK 500
QY 541 SKKGKMKROYSLSSSEBELASTPEYTSODVLESESVSEKGDOKGRKRTSEOGVLSDS 600
Db 501 SKKGKMKROYSLSSSEBELASTPEYTSODVLESESVSEKGDOME----- 545
QY 601 NTFSEROKKMYGGHSLLEDLEWSEPOIKDSGYDTGCSSTLNEHSHSDKHPTWOPSK 660
Db 546 -----TSMLEHSAWH-----SSASPMSLHPVTWOPSK 573
QY 661 DGDRLIGRILNLRKLDGSPYRDSGAMGLKVGGKMTESGRLCATFKVKKGLADYVG 720
Db 574 DGDRLIGRILNLRKLDGSPYRDSGAMGLKVGGKMTESGRLCATFKVKKGLADYVG 633
QY 721 HLRPGDEVLWNGRLLQGAFFEEYNNILLESKPEPQVELVSRPIGDIPRLPDSTHAOLE 780
Db 634 HLRPGDEVLWNGRLLQGAFFEEYNNILLESKPEPQVELVSRPIGDIPRLPDSTHAOLE 693
QY 781 SSSSPESOKMDRPSISVTPSPSGMLRDVPOFLSGOLSTLMPKVKHOLIVTTLGAKD 840
Db 694 SSSSPESOKMDRPSISVTPSPSGMLRDVPOFLSGOLSTLMPKVKHOLIVTTLGAKD 753
QY 841 LPSREDGRPNRPVYKTYFLPDRSDKNRKRTKYVKTLEPKNNQFTIYSPVHREPERML 900
Db 754 LPSREDGRPNRPVYKTYFLPDRSDKNRKRTKYVKTLEPKNNQFTIYSPVHREPERML 813
QY 901 EITLMQOARKRESESEFLGELILELFTALLDDEPHWYKLQTHDVSSLPRLPPSPYLPRRQ 960
Db 814 EITLMQOARKRESESEFLGELILELFTALLDDEPHWYKLQTHDVSSLPRLPPSPYLPRRQ 873
QY 961 LHGESPTRLQORSKRISDSSEVDYDCEDGVGVSDYRHNGRDLQSTLSVPEQVWSSNHC 1020
Db 874 LHGESPTRLQORSKRISDSSEVDYDCEDGVGVSDYRHNGRDLQSTLSVPEQVWSSNHC 933
QY 1021 SPGSGPHRVDVIGRTSRMSPSAPPORNVEGHRGTRATGHTNTISRMDRHRVMDHYS 1080
Db 934 SPGSGPHRVDVIGRTSRMSPSAPPORNVEGHRGTRATGHTNTISRMDRHRVMDHYS 993
QY 1081 DRD-----RDCEAADROPYHRSRSTEOPLLERTTTRSSSER 1118
Db 994 ERDSHFLLTPRSRHRQTSSEHHNRDGRDCEAADROPYHRSRSTEOPLLERTTTRSSSER 1053
QY 1119 PPTNLMRSPSLMTGSRAPSPALSRSHPRGSAVOTSSSPGGRGRGROLPOLPPKCTL 1178
Db 1054 ADTNLMKSMPSLMTGSRAPSPALSRSHPRGSAVOTSSSPGGRGRGROLPOLPPKCTL 1113
QY 1179 ER-----SAMDIEERNQMLINKYQVAGSDPRLQDYSKTRSGMDPHRG 1224
Db 1114 EMTIEDMDSTRKRNAGAMDIEERNQMLINKYQVAGSDPRLQDYSKTRSGMDPHRG 1173
QY 1225 ADTVSTKSSSDVSDVSAVSRSTSASRFSSTIYSVOSERPGRNKKISVFTSKMNRQMG 1284
Db 1174 ADTVSTKSSSDVSDVSAVSRSTSASRFSSTIYSVOSERPGRNKKISVFTSKMNRQMG 1233
```

```
QY 1285 VSGKNLTKSTISGDMCSLEKNDSQSPTAVGALGTSCKKKRRSSTGAKMVAIVGLSRKS 1344
Db 1234 VSGKSMKSTISGDMCSLEKNDSQSPTAVGALGTSCKKKRRSSTGAKMVAIVGLSRKS 1293
QY 1345 SASOLQTEGGGKKLRSTVQSTETGLAVERNMMTRQASREESTGSKNSYSSEGNLFFP 1404
Db 1294 SASOLQTEGGGKKLRSTVQSTETGLAVERNMMTRQASREESTGSKNSYSSEGNLFFP 1353
QY 1405 GVRLASDSQSFIDLGLGPAOLVGRQTLATPAMGDIOYGMMDKQOLEVEITRAGLYVK 1464
Db 1354 GVRLASDSQSFIDLGLGPAOLVGRQTLATPAMGDIOYGMMDKQOLEVEITRAGLYVK 1413
QY 1465 PGSKTLPPAPYKYYLLDNGVCIAKKTKRVAKRTLEPLYQOOLSEEPSQGRVLOIIVMGD 1524
Db 1414 PGSKTLPPAPYKYYLLDNGVCIAKKTKRVAKRTLEPLYQOOLSEEPSQGRVLOIIVMGD 1473
QY 1525 YGRMDHKSFGVQAQIILDELELSNMVIGWFKLFPSSLVDP TSAPLTRRASQSSLESSTG 1584
Db 1474 YGRMDHKSFGVQAQIILDELELSNMVIGWFKLFPSSLVDP TSAPLTRRASQSSLESSTG 1533
QY 1585 PYSYRS 1590
Db 1534 PYSYRS 1539
RESULT 4
QY 09JIS1 PRELIMINARY; PRT; 1555 AA.
ID 09JIS1;
AC 09JIS1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE RIM2.
GN RIM2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347919; PubMed=10748113;
RA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins. Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins.";
RL J. Biol. Chem. 275:20033-20044(2000).
DR EMBL; AF199322; AAF81644.1; -.
DR InterPro; IPR000008; -.
DR InterPro; IPR000306; -.
DR InterPro; IPR000345; -.
DR InterPro; IPR001478; -.
DR Pfam; PF00168; C2; 2.
DR Pfam; PF00595; PDZ; 1.
DR PROSITE; PS50004; C2_DOMAIN_2; 2.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART; SM00239; C2; 1.
SQ SEQUENCE 1555 AA; 175912 MW; D76967BB36D9704E CRC64;
```

Query Match 89.6%; Score 7420.5; DB 11; Length 1555;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 1451; Conservative 21; Mismatches 31; Indels 139; Gaps 6;

```
QY 1 MSAPLGRGRPAPTPAASQPPPEMPDLSHLTBEERKIIIAVMDROKKEEKEQSVLKI 60
Db 1 MSAPLGRGRPAPTPAASQPPPEMPDLSHLTBEERKIIIAVMDROKKEEKEQSVLKI 59
QY 61 KEHKAQPTQWPFPSGTTELNVNVLQPOQKPNKEPQTKLHQQFEMKQGVKKMGESQ 120
Db 60 -----KLHQQFEMKQGVKKMGESQ 80
QY 121 QOQEOKGAFTCGICHTKPADGCHNCSCQTFKCARCGGAVLSRNSKVMVNCMLCRQ 180
Db 121 QOQEOKGAFTCGICHTKPADGCHNCSCQTFKCARCGGAVLSRNSKVMVNCMLCRQ 180
```



```

Db 81 QOQOBKGDAPTGCIGCHTKFADGCGHNCSCQTKFCARCGRVSLRKNKVMVYCNLCRKQ 140
QY 181 QELLTKSGAMFYNSGNTLQOPDOKVPRGLRMEAPQOEKKAKLHEQPOFOGARGDLSPVA 240
Db 141 QELLTKSGAMFYNSGNTLQOPDOKALRKLSEAPQOEKKAKLHEQTOFGPGDSSVPA 200
QY 241 VEKGRAGILTRQDTTKNGSGVHQIASDMPSDRKRSRPSVSDQNRREYDSEEREDYSQY 300
Db 201 VERGRAGILTRQDSTKNGSGMKHQIASDMPSDRKRSRPSVSDQNRREYDSEEREDYSQY 260
QY 301 PSDGTMPSPSDYADRSRQREPOFYEEPGHLYRDSNRGRHRSKEYIVDEDEVEDSREX 360
Db 261 PSDGTMPSPSDYADRSRQREPOFYEEPGHLYRDSNRGRHRSKEYIVDEDEVEDSREX 320
QY 361 ERORREBEYQARYSDPRLAYRVPKPOPYEEQMRTHAEVSRARHERHRSVSLANAELED 420
Db 321 ERORREBEYQARYSDPRLAYRVPKPOPYEEQMRTHAEVSRARHERHRSVSLANAELED 380
QY 421 SRISLLRMDRPSRQSVSEBRRAAMENORSYMERTRRAGOSSTPORTSNHSPTRRSP 480
Db 381 SRISLLRMDRPSRQSVSEBRRAAMENORSYMERTRRAGOSSTPORTSNHSPTRRSP 440
QY 481 IPLDRPMRRADSLRKHHLDPSSAVRKTREKMETMLRNDLSLSDQSESVRPPRPK 540
Db 441 IPLDRPELRADSLRKHHLDPSSAVRKTREKMETMLRNDLSLSDQSESVRPPRPK 500
QY 541 SKKGKMKQVSLSSSEELASTPEYTCDDVLESESVESEKDGSKKRTSQGVLSDS 600
Db 501 SKKGKMKQVSLSSSEELASTPEYTCDDVLESESVESEKDGME----- 545
QY 601 NTRSERQKRMYYGHSLEEDLEMSERPOLKDGSDVTCSTLLNEHSHSDKHPTWQPSK 660
Db 546 -----YSMLHSAHWH-----SSEASPMKSLHVTWQPSK 573
QY 661 DGDRLIGRILLNKLKDGSVPRDSCAMGLKVGCGKMTESGRICAFITTKYKKSGLADTVG 720
Db 574 DGDRLIGRILLNKLKDGSVPRDSCAMGLKVGCGKMTESGRICAFITTKYKKSGLADTVG 633
QY 721 HLRGDEVLENNGRLLQCATPEEYVNIILLESKPEQVELVNSRPIDIRIPSTHAOLE 780
Db 634 HLRGDEVLENNGRLLQCATPEEYVNIILLESKPEQVELVNSRPIDIRIPSTHAOLE 693
QY 781 SSSSFESQKMDRPSISVTSKPMGLRDVPOFLSGOLS-----IKLWF 824
Db 694 SSSSFESQKMDRPSISVTSKPMGLRDVPOFLSGOLS-----IKLWF 753
QY 825 DKVGHQLIVTLLGAKDLPSREDGRPRNPVYKIFLPRDSDKNKRRTKYVKKLEPKWQNT 884
Db 754 DKVGHQLIVTLLGAKDLPSREDGRPRNPVYKIFLPRDSDKNKRRTKYVKKLEPKWQNT 813
QY 885 FIYSVPHRERERLETTIMDOARVRESESEFLGILTELEFALLDDDEPHWKLOTHDV 944
Db 814 FIYSVPHRERERLETTIMDOARVRESESEFLGILTELEFALLDDDEPHWKLOTHDV 873
QY 945 SSELPLRPSPYLPRRQLGSEPTRRQLQSKRISLSEVSDYDCEDGCVVSDYRHNGRDLQ 1004
Db 874 SSELPLRPSPYLPRRQLGSEPTRRQLQSKRISLSEVSDYDCEDGCVVSDYRHNGRDLQ 933
QY 1005 SSTLSVPEQVWSSNHCSPSGSHRVDVIGRTKRSWSPSAPPPORANVOGHGRTATGTYNT 1064
Db 934 SSTLSVPEQVWSSNHCSPSGSHRVDVIGRTKRSWSPSAPPPORANVOGHGRTATGTYNT 993
QY 1065 ISRMDRHRVMDHYSSDD-----RCEAADRPPYHRSRTEQ 1102
Db 994 ISRMDRHRVMDHYSSDDSHFLTLPRSRHROTSEHNRDGRCEAADRPPYHRSRTEQ 1053
QY 1103 RPLLERTTSSRSEPRPTNLMRSMPSLMTGSRAPSPALSRSHPRTGSVOTSPSSPTGT 1162
Db 1054 RPLLERTTSSRSEPRPTNLMRSMPSLMTGSRAPSPALSRSHPRTGSVOTSPSSPTGT 1113
QY 1163 GRGRQOLPOLPPKGLER-----SANDIEBRNROMKLNKYQVAGSDPRLE 1208
Db 1114 GRGRQOLPOLPPKGLERITEDMDSTRKNSGAMDIEBRNROMKLNKYQVAGSDPRLE 1173

```

```

QY 1209 ODYHSKYSRGMDFHRGADTVSTRKSSDSVDVSAVSRSTSSASRFSSTYSKVSOSERPRGN 1268
Db 1174 ODYHSKYSRGMDFHRGADTVSTRKSSDSVDVSAVSRSTSSASRFSSTYSKVSOSERPRGN 1233
QY 1269 KRISVETSKMONRQMGVSGKNTLKTSTISGDMCSLEKNDGSDTDVAGALGTSGKKRRSS 1328
Db 1224 KRISVETSKMONRQMGVSGKNTLKTSTISGDMCSLEKNDGSDTDVAGALGTSGKKRRSS 1293
QY 1329 IGAKVAIVGLSRKSRASQLSQTEBGGKKLASTVORSTETGLAVEMRNMTRQASREST 1388
Db 1294 IGAKVAIVGLSRKSRASQLSQTEBGGKKLASTVORSTETGLAVEMRNMTRQASREST 1353
QY 1389 DGSNMSYSEGNLIFPGVVALDSQSFLLDGLGPAQOLVGRQTLAPAMGDIOVGMMDK 1448
Db 1354 DGSNMSYSEGNLIFPGVVALDSQSFLLDGLGPAQOLVGRQTLAPAMGDIOVGMMDK 1413
QY 1449 GQLEVEIIRARGLVYKPGSKTLPAPYVKKYLLDNGVCIAKKTKVARKTLEPLYOQLLSE 1508
Db 1414 GQLEVEIIRARGLVYKPGSKTLPAPYVKKYLLDNGVCIAKKTKVARKTLEPLYOQLLSE 1473
QY 1509 EESPOGRVLIIVMGDYGMDHKSFEWGAQILDELELSNMVIGWEKLEPPSSLVDPTSA 1568
Db 1474 EESPOGRVLIIVMGDYGMDHKSFEWGAQILDELELSNMVIGWEKLEPPSSLVDPTSA 1533
QY 1569 PLTRRASOSLESSTGSPYSRS 1590
Db 1534 PLTRRASOSLESSTGSPYSRS 1555

RESULT 5
09JUISO PRELIMINARY: PRT: 1399 AA.
AC 09JUISO:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RIM2-2B.
GN RIM2-2B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347919; PubMed=10748113;
RA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins: Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins.";
RL J. Biol. Chem. 275:20033-20044(2000).
DR EMBL; AF199324; AAF81646.1; -.
DR InterPro; IPR000008; -.
DR InterPro; IPR000306; -.
DR InterPro; IPR000345; -.
DR InterPro; IPR001478; -.
DR Pfam; PF00168; C2; 2.
DR Pfam; PF00595; PDZ; 1.
DR PROSITE; PS50004; C2 DOMAIN_2; 2.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART; SM00239; C2; 1.
SQ SEQUENCE 1399 AA; 158945 MW; 3F1550E3D52D237B CRC64;

Query Match 82.8%; Score 6856.5; DB 11; Length 1399;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1338; Conservative 17; Mismatches 22; Indels 235; Gaps 3;

QY 1 MSAPLGPGRAPPTPAASOPPOPEMDLSHLEERKIIIVAMDROKKEEKOQVYKI 60
Db 1 MSAPLGPGRAPPTPAASOPPOPEMDLSHLEERKIIIVAMDROKKEEKOQVYKI 59
QY 61 KEHKAQPTOWPFSGITELVNVNVLQPOOKOPNEKEPOTKLHOQFEYKQVYKMGESQ 120
Db 1 KEHKAQPTOWPFSGITELVNVNVLQPOOKOPNEKEPOTKLHOQFEYKQVYKMGESQ 119

```

```
Db 60 -----KLHQFEMYKEQVKMGESQ 80
QY 121 QOQEQKADPTGICHTKTKFADGGCHNCYCQTKPCARCGRVLSRSKKVMVAVCULCRQ 180
Db 81 QOQEQKADPTGICHTKTKFADGGCHNCYCQTKPCARCGRVLSRSKKVMVAVCULCRQ 140
QY 181 QELITKSGAMFYNSGNTLQOPDQKVPGLNEBAPQEKKKAKLHEQPOFOGAGDLSVPA 240
Db 141 QELITKSGAMFYNSGNTLQOPDQKVPGLNEBAPQEKKKAKLHEQPOFOGAGDLSVPA 200
QY 241 VEGRAHAGLRTOTIKNGSGVKKHDIASDMPDRKRSVSDONRKYQSEEREDYSQV 300
Db 201 VEGRAHAGLRTOTIKNGSGVKKHDIASDMPDRKRSVSDONRKYQSEEREDYSQV 260
QY 301 PSQGTMRPSDADRRSOREPOPEYEEGHLYNDRSDNRGRHSHKEYVDDQEDVSREX 360
Db 261 PSQGTMRPSDADRRSOREPOPEYEEGHLYNDRSDNRGRHSHKEYVDDQEDVSREX 320
QY 361 ERQREEEYQARYRSDPNLARVPYKPOPEYEQMRHAEVSRARERHSDVSLANAELED 420
Db 321 ERQREEEYQARYRSDPNLARVPYKPOPEYEQMRHAEVSRARERHSDVSLANAELED 380
QY 421 SRTSLRMBRPSQORSYSEERRAMENORSTSMERTREAGOSSTPQRTSNHSPTPRSP 480
Db 381 SRTSLRMBRPSQORSYSEERRAMENORSTSMERTREAGOSSTPQRTSNHSPTPRSP 440
QY 481 IPDRPDMRRADSLRKOHHLDPSSAVAKTKREKMETMLRNDLSLSDOSESVRPPRRHK 540
Db 441 IPDRPDMRRADSLRKOHHLDPSSAVAKTKREKMETMLRNDLSLSDOSESVRPPRRHK 500
QY 541 SKKGAMROVSLSSSEELASTPEYTSQDDVELESESVSEKGSOKGRKRTSEQVLSDS 600
Db 501 SKKGAMROVSLSSSEELASTPEYTSQDDVELESESVSEKGSOKGRKRTSEQVLSDS 560
QY 601 NTRSEOKRMVYGGSGSLSEDLSEMSPEQIKDSQVDTSCSTTLNEBHSKDHPTVWQSK 660
Db 561 NTRSEOKRMVYGGSGSLSEDLSEMSPEQIKDSQVDTSCSTTLNEBHSKDHPTVWQSK 620
QY 661 DDDRILGRILNKLKLDGSPYRDSGAMLGKLVGKMTESGRLCFAFTTKVKGSLADTVG 720
Db 621 DDDRILGRILNKLKLDGSPYRDSGAMLGKLVGKMTESGRLCFAFTTKVKGSLADTVG 680
QY 721 HLRPGDEVLEWNGRLIQGATFEEVYNIILSKRPEQVELVSRPDIGDIPRIIPDSTHAOLE 780
Db 681 HLRPGDEVLEWNGRLIQGATFEEVYNIILSKRPEQVELVSRPDIGDIPRIIPDSTHAOLE 740
QY 781 SSSSSSEOKMRRPSISVSPMSPGMLRDVPOFLSGOLSTIKLMDKXGHOLIVTILKAD 840
Db 741 SSSSSSEOKMRRPSISVSPMSPGMLRDVPOFLSGOLSTIKLMDKXGHOLIVTILKAD 800
QY 841 LPSREDEGRPNPVKITYFLPDRSDKNKRTKTYKKTLEPKWNOFTIYSPVHRREFRERML 900
Db 801 LPSREDEGRPNPVKITYFLPDRSDKNKRTKTYKKTLEPKWNOFTIYSPVHRREFRERML 860
QY 901 EITLMDQANVRESESEFLGEIILELTALLDDEPHWYKLTQHDVSSDLPRPSPYLPKRQ 960
Db 861 EITLMDQANVRESESEFLGEIILELTALLDDEPHWYKLTQHDVSSDLPRPSPYLPKRQ 920
QY 961 LHGESPTRLQSKRISDSEVSYDCEDEGVGVSDYKHNHGRDLOQSTLSLSPVEQVSSNHC 1020
Db 921 LHGESPTRLQSKRISDSEVSYDCEDEGVGVSDYKHNHGRDLOQSTLSLSPVEQVSSNHC 980
QY 1021 SPSSGPHRDVIGRTSRMSPSPAPPOANVBOGHRGTRATGHYNTISMDDRRHVMDDHYS 1080
Db 981 SPSSGPHRDVIGRTSRMSPSPAPPOANVBOGHRGTRATGHYNTISMDDRRHVMDDHYS 1040
QY 1081 DRD-----RDCEAADQRYHRSHTSEQRPPLLERTTTRSRSR 1118
Db 1041 EEDSHFLTLPRSRHROTSEHHNHRDCEAADRQYHRSHTSEQRPPLLERTTTRSRSR 1100
QY 1119 PDTNLMRSPSLMTGSGAPRSPALSSHRTRTGSVQTSPPSTPGTGRGRQOLPOLPRPGTL 1178
Db 1101 ADTNLMRSPSLMTGSGAPRSPALSSHRTRTGSVQTSPPSTPGTGRGRQOLPOLPRPGTL 1160
```

```
QY 1179 ERSAMDIEERNROMKLNKYQVAGSDPRLEDYHASKTRSGMDPHRADTVSTKSSDPSV 1238
Db 1161 ER----- 1162
QY 1239 DVSAVSRSTSASRSTSTSYMVSQSERPRGNRKISVTSKMNQNRQMGVSGKMLTSTISG 1298
Db 1163 ----- 1162
QY 1299 DMCLEKNDGSQSDTAVAGALGTSGKKRRSSIGAKMAIVGLSRKRSASOLSQTEGGGKK 1358
Db 1163 -----GGGKK 1167
QY 1359 LRSTVORSTETGLAVERNMNTQASRSTSDGSNYSSEBGNLFFPGVRLASDSQSPFL 1418
Db 1168 LRSTVORSTETGLAVERNMNTQASRSTSDGSNYSSEBGNLFFPGVRLASDSQSPFL 1227
QY 1419 DGLGPAQLVGRQTLATPAMGDIOVMDDKKGQLEVEITIRAGLVVKKPGSKTLPAIPYKVY 1478
Db 1228 DGLGPAQLVGRQTLATPAMGDIOVMDDKKGQLEVEITIRAGLVVKKPGSKTLPAIPYKVY 1287
QY 1479 LLDNGVCIAKKKTIVAKTLEPLYQOLLSEESPOGRVLOIYWGDIYGRMDHKSFMGVAQ 1538
Db 1288 LLDNGVCIAKKKTIVAKTLEPLYQOLLSEESPOGRVLOIYWGDIYGRMDHKSFMGVAQ 1347
QY 1539 ILDELELSNMVTGWFKLPPSSLVDPSTAPLITRRASOSSLESSTGSPYSRS 1590
Db 1348 ILDELELSNMVTGWFKLPPSSLVDPSTAPLITRRASOSSLESSTGSPYSRS 1399
```

```
RESULT 6
09JIR7
AC 09JIR7 PRELIMINARY; PRT: 1330 AA.
DT 01-OCT-2000 (TREMBLrel, 15, Created)
DT 01-OCT-2000 (TREMBLrel, 15, last sequence update)
DT 01-MAR-2001 (TREMBLrel, 16, last annotation update)
DE RIM2-4C.
GN RIM2-4C.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=203447919; Pubmed=10748113;
RA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins. Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins."
RL J. Biol. Chem. 275:20033-20044(2000).
DR EMBL: AF199329; AAF81651.1; -.
DR InterPro: IPR000008; -.
DR InterPro: IPR000306; -.
DR InterPro: IPR000345; -.
DR InterPro: IPR001478; -.
DR Pfam: PF00168; C2: 2.
DR Pfam: PF00595; PDZ: 1.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART: SM00239; C2: 1.
SQ SEQUENCE 1330 AA; 151044 MW; 485906E1F66FC73 CRC64;
```

```
Query Match 79.0%; Score 6543; DB 11; Length 1330;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 1283; Conservative 18; Mismatches 29; Indels 260; Gaps 4;
QY 1 MSAPLPGRCRPAPTPAASOPPOPEMDLSHLEERKIIILAVMDQKKEEKGQSVLKI 60
Db 1 MSAPLPGRCRPAPTPAASOPPOPEMDLSHLEERKIIILAVMDQKKEEKGQSVLKI 59
QY 61 KEENKAOPTOWMPFSGITELVANNVLIPOOKOPNEKEPOTKILHQFEMYKEQVKMGESQ 120
Db 1 KEENKAOPTOWMPFSGITELVANNVLIPOOKOPNEKEPOTKILHQFEMYKEQVKMGESQ 119
```

```

Db 60 ----- KLHQEFENYKBEQVKKMGESQ 80
QY 121 QOOROKGAPTCIGCHKTKEADGCGHNCSCOTKPCARCGGRVSLRKNKMYCNLCRKQ 180
Db 81 QOOROKGAPTCIGCHKTKEADGCGHNCSCOTKPCARCGGRVSLRKNKMYCNLCRKQ 140
QY 181 QEILTKSGAMFYNSGNTLQOPDQKVPRLRNEBEARPOEKKAKLHEOPQFOGAGDLSVPA 240
Db 141 QEILTKSGAMFYNSGNTLQOPDQKALRGLRSEBEARPOEKKAKLHEOTQFOGPPGDSSVPA 200
QY 241 VEKGRAGLTPDQITKNGSVKQIADMDSPDKRSPSVSDONRRYEOSEEREDYSQYV 300
Db 201 VERGRAGLTPDQITKNGSVKQIADMDSPDKRSPSVSDONRRYEOSEEREDYSQYV 260
QY 301 PSDGTMPSPDYADRRSOREPOFYEEPHLNYRDSNRGRHRSKETYVDEDEVESDEX 360
Db 261 PSDGTMPSPDYADRRSOREPOFYEEPHLNYRDSNRGRHRSKETYVDEDEVESDEX 320
QY 361 ERORREERYQARYRSDPNLARYPVKPOYEEOMRIHAEVSRAHERHSDVSLAMELED 420
Db 321 ERORREERYQARYRSDPNLARYPVKPOYEEOMRIHAEVSRAHERHSDVSLAMELED 380
QY 421 SRISLIRMDRPSRQSVSEBRAMENORSYSMEKTRERAGOSSTYPORTNHSPTPRSP 480
Db 381 SRISLIRMDRPSRQSVSEBRAMENORSYSMEKTRERAGOSSTYPORTNHSPTPRSP 440
QY 481 IPLRDPKRRADSLRKHHLDPSSAVRKTREKMETMLRNDLSLSSOSESVRPPPHK 540
Db 441 IPLRDPKRRADSLRKHHLDPSSAVRKTREKMETMLRNDLSLSSOSESVRPPPHK 500
QY 541 SKKGKMRQVSLSSSEELASTPEYTSQDVELESSEVSEKDGKRTSEGVLSDS 600
Db 501 SKKGKMRQVSLSSSEELASTPEYTSQDVELESSEVSEKDGKRTSEGVLSDS 560
QY 601 NTRSEROKKRMVYGGHSLSEEDLEKSEPOITKDSQVDTCSSTTLNEBHSKDKHPVYWPSPK 660
Db 560 NTRSEROKKRMVYGGHSLSEEDLEKSEPOITKDSQVDTCSSTTLNEBHSKDKHPVYWPSPK 620
QY 661 DGDRILGRILNLRKLDKSVPRDSGMLGKLVGCKMTESGRCAPFTYKKKSLADTVG 720
Db 720 DGDRILGRILNLRKLDKSVPRDSGMLGKLVGCKMTESGRCAPFTYKKKSLADTVG 680
QY 721 HLRPDEVLEWNGRLLOGATFEVYNIILSEKPEPOVELVSRPIGDIRPIPOSTHAOLE 780
Db 680 HLRPDEVLEWNGRLLOGATFEVYNIILSEKPEPOVELVSRPIGDIRPIPOSTHAOLE 740
QY 781 SSSSFESQKMDRPSISVYSPMSPGMLRDVPOFLSGQLSIKLMFDEVGHOLIYTLIGAKD 840
Db 740 SSSSFESQKMDRPSISVYSPMSPGMLRDVPOFLSGQLSIKLMFDEVGHOLIYTLIGAKD 800
QY 841 LPSRDEGPRRNPVYIYFLPDRSDKNKRTKTVKTLEPKMNQTFIYSPVHRREFRRL 900
Db 900 LPSRDEGPRRNPVYIYFLPDRSDKNKRTKTVKTLEPKMNQTFIYSPVHRREFRRL 860
QY 901 EITLMDQARVRESESEFGEIILIEETALLDEPHWYKLTQTHVSSLPRLRSPYLLRQ 960
Db 860 EITLMDQARVRESESEFGEIILIEETALLDEPHWYKLTQTHVSSLPRLRSPYLLRQ 920
QY 961 LHGESPTRLQSRKISDSSEVSDYDCESGVYSDYRHNGRDQSSSTLSVPEOVMSSNHC 1020
Db 874 LHGESPTRLQSRKISDSSEVSDYDCESGVYSDYRHNGRDQSSSTLSVPEOVMSSNHC 933
QY 1021 SPSSGPHRVNDYIGTRRSPSPAPPOARNVEOGRGRATGHYNTISMDHRRVMDHYSS 1080
Db 934 SPSSGPHRVNDYIGTRRSPSPAPPOARNVEOGRGRATGHYNTISMDHRRVMDHYSS 993
QY 1081 DRDRDCEAADQPYHRSHSTEGRPLLERTTTSRSSSERPDTNLMRSMPLMTGRSAPSP 1140
Db 994 ERDRDCEAADQPYHRSHSTEGRPLLERTTTSRSSSERPDTNLMRSMPLMTGRSAPSP 1053
QY 1141 ALSRSHPTGVSQVSPSSSTPGGRGRGROLPOLPKGTLERSAMDIERNNQMLNKTKQY 1200
Db 1054 ALSRSHPTGVSQVSPSSSTPGGRGRGROLPOLPKGTLERSAMDIERNNQMLNKTKQY 1093

```

```

QY 1201 AGSDPRLBDYHSKYRSGMDPHRGADTVSTKSSSDSDVSDVSAVSRTSSASRFSSTYSMV 1260
Db 1094 ----- 1093
QY 1261 QSERPRGNKISVFTSKMONRQMGVSGKMLTKSTISISGDMCSLEKNDGSDTRAVGALGT 1320
Db 1094 ----- 1093
QY 1321 SKKRRSSIGAMVAIVGLSRKRSRASAQSQTDEGGKKLRSTYQORSTETGLAEMRWMT 1380
Db 1094 ----- GGGKKLRSTYQORSTETGLAEMRWMT 1120
QY 1381 ROASRESTDGSNNYSSEGNLIFPGVRLASDSQSFPLDGLGPAQLGROTLPAMGDI 1440
Db 1121 ROASRESTDGSNNYSSEGNLIFPGVRLASDSQSFPLDGLGPAQLGROTLPAMGDI 1180
QY 1441 QVGMDDKKGQLEVEIIRARGLVKPGSKTLPAPYVYLLDNGVCIAKKKTKVARKTLEP 1500
Db 1181 QVGMDDKKGQLEVEIIRARGLVKPGSKTLPAPYVYLLDNGVCIAKKKTKVARKTLEP 1240
QY 1501 LYQQLSFEESPOGRVLOITWGDYGRMDKSPMGVAQIILDELESNNYIGFKLEPPS 1560
Db 1241 LYQQLSFEESPOGRVLOITWGDYGRMDKSPMGVAQIILDELESNNYIGFKLEPPS 1300
QY 1561 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1590
Db 1301 SLVDPTSAPLTRRASQSSLESSTGSPYSRS 1330

```

```

RESULT 7
Q9JHU6 PRELIMINARY; PRT: 1352 AA.
ID Q9JHU6
AC Q9JHU6
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE RIM2-5C (RIM2-2A) (RIM2-3B) (RIM2-4A).
GN RIM2-5C OR RIM2-2A OR RIM2-3B OR RIM2-4A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347919; PubMed=10748113;
KA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/NIM Family of Neuronal C2 Domain Proteins. Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins.";
RL J. Biol. Chem. 275:20034-20044(2000).
DR EMBL: AF199332; AAF81654.1; -
DR EMBL: AF199326; AAF81648.1; -
DR EMBL: AF199327; AAF81649.1; -
DR InterPro: IPR000008; -
DR InterPro: IPR000306; -
DR InterPro: IPR000345; -
DR InterPro: IPR001478; -
DR Pfam: PF00168; C2_2; -
DR Pfam: PF00595; PDZ_1; -
DR PROSITE: PS50004; C2-DOMAIN_2; 2.
DR SMART: SM00239; C2_1; -
DR SMART: SM00239; C2_1; -
SQ SEQUENCE 1352 AA; 153718 MW; 5E1E6F0E42DFEA9C CRC64;

```

```

Query Match 78.7%; Score 6522; DB 11; Length 1352;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 1283; Conservative 18; Mismatches 29; Indels 282; Gaps 5;
QY 1 MSAPLGRPRAPVTPAASQPPQPEMPDLSHLLEERKIIILAVMDRKKKEEKQSVLKI 60
Db 1 MSAPLGRPRAPVTPAASQPPQPEMPDLSHLLEERKIIILAVMDRKKKEEKQSVLKI 59

```

```

QY 61 KEHKAQPTQWFPESGITELVNNVLQPOQKQPNKEKQPOTKLHQEFEMTKQEQYKKMGESQ 120
Db -----KTHQOEFEMTKQEQYKKMGESQ 80
QY 121 QOQOBOKDAPTQICGICHTKFKADGGCGHNCYCQTKFCARGCGVSLRSNKVMVMVNCILCRQ 180
Db 81 QOQOBOKDAPTQICGICHTKFKADGGCGHNCYCQTKFCARGCGVSLRSNKVMVMVNCILCRQ 140
QY 181 QEILTKGAFYNGSNTLQOPDQKVPFGLNENAPQEKKAHLHQPQFQAGDLISVA 240
Db 141 QEILTKGAFYNGSNTLQOPDQKVPFGLNENAPQEKKAHLHQPQFQAGDLISVA 200
QY 241 VEGRAGLITROPTIKNGSGVYKHQIASDMPDRKSPSVSDONRNYQSEEREDYSQV 300
Db 201 VEGRAGLITROPTIKNGSGVYKHQIASDMPDRKSPSVSDONRNYQSEEREDYSQV 260
QY 301 PSDGTMRSPSDYADRRSOPREPOFEERPHLNYRDSNRGRHSKEYIYDDEDVESRDEY 360
Db 261 PSDGTMRSPSDYADRRSOPREPOFEERPHLNYRDSNRGRHSKEYIYDDEDVESRDEY 320
QY 361 ERORREERYARISDNLAIRPYKPOPEEQMKRIHAESRARRHRRSHDVSILANAELED 420
Db 321 ERORREERYARISDNLAIRPYKPOPEEQMKRIHAESRARRHRRSHDVSILANAELED 380
QY 421 SRTSLRLMDPRRSOVSSEERRAAMENORSYSMERTREAOGOSYPORTNSIPTPRRSP 480
Db 381 SRTSLRLMDPRRSOVSSEERRAAMENORSYSMERTREAOGOSYPORTNSIPTPRRSP 440
QY 481 IPLDRPDMRRADSLRKQHHLDPSSAVRKTREKMETMLRNDLSLSDQSESVRPPRRHK 540
Db 441 IPLDRPDMRRADSLRKQHHLDPSSAVRKTREKMETMLRNDLSLSDQSESVRPPRRHK 500
QY 541 SKKGKMKROYLSLSSSEBELASTPEYTSQDVELESESEKGDQCKRKRTSBOGVLSDS 600
Db 501 SKKGKMKROYLSLSSSEBELASTPEYTSQDVELESESEKGDQCKRKRTSBOGVLSDS 545
QY 601 NTRSEBOKRMYYGSHLEEDLEWSEPOIKDSGYDTCSTTLNEHSHSDKHPYTWQPSK 660
Db 546 -----YSMLHASMWH-----SSEASMSLHPYTWQPSK 573
QY 661 DGDRLILGRILLNRLKLDGVSPPRSGAMLGLKVGKMTESGRICAFITTKVKGSLADTVG 720
Db 574 DGDRLILGRILLNRLKLDGVSPPRSGAMLGLKVGKMTESGRICAFITTKVKGSLADTVG 633
QY 721 HLRPGDEVLEWNGRLLOGATFEERYNTILLESKPEPOVEIYVSRIGDIPRIPDSTHGLE 780
Db 634 HLRPGDEVLEWNGRLLOGATFEERYNTILLESKPEPOVEIYVSRIGDIPRIPDSTHGLE 693
QY 781 SSSSFSSESQKMDRPSISVTSPPMSPGMLRDVPOPLSGQLSIKLMPDKYGHQIIVITLIGAKD 840
Db 694 SSSSFSSESQKMDRPSISVTSPPMSPGMLRDVPOPLSGQLSIKLMPDKYGHQIIVITLIGAKD 753
QY 841 LBSREGRRRNPVYKITYFLPDRSDKNNKRTKYKTKLEPKWNQTFIYSPVHRRFEREML 900
Db 754 LBSREGRRRNPVYKITYFLPDRSDKNNKRTKYKTKLEPKWNQTFIYSPVHRRFEREML 813
QY 901 ETLIMQOARVRESESEFLGFIETLEETALLDDPWHYKLTQTHOVSSILPRLPSPSYLPRQ 960
Db 814 ETLIMQOARVRESESEFLGFIETLEETALLDDPWHYKLTQTHOVSSILPRLPSPSYLPRQ 873
QY 961 LHGESFTRILQSRKISLSEVSDYDCEDEGVGVSDYRHHNGRDLQSSLTSLVEQVMSSNHC 1020
Db 874 LHGESFTRILQSRKISLSEVSDYDCEDEGVGVSDYRHHNGRDLQSSLTSLVEQVMSSNHC 933
QY 1021 SPSSGSHRDVIGRTSWSBSPAPRPNANVEQGRGTRATGHYNTISMDRHRVMDHYSS 1080
Db 934 SPSSGSHRDVIGRTSWSBSPAPRPNANVEQGRGTRATGHYNTISMDRHRVMDHYSS 993
QY 1081 DRD-----RDCEAADROPYHRSNRSTORPLERTTTRSRSSER 1118
Db 994 ERDHSFLLTLPBSRKHROTSEHHNRDGDCCRAADROPYHRSNRSTORPLERTTTRSRSSER 1053

```

```

QY 1119 PDTNLKMSMPSLMTGNSAPPSPALSRSHPRGTGSVQTSBSSPTGTGRGRQDLPQLEPKCTL 1178
Db 1054 ADYNLMKMSMPSLMTGNSAPPSPALSRSHPRGTGSVQTSBSSPTGTGRGRQDLPQLEPKCTL 1113
QY 1179 ERSAMDIEERNRQMKLNKYQVAGSDPRLBDYHSXKRSRGMDPRRGADVTSTKSDSVS 1238
Db 1114 ER----- 1115
QY 1239 DVSASRTSSASRFSSTSYMSVOSERPRGNRKISVFTSKMQRQMGVSGKNLTXTSTISG 1298
Db 1116 ----- 1115
QY 1299 DMSLEKNDGOSQDVAVALGCTSGKKRRSSITGAKMAIIVGLSKRSASQLSQTEGGGKK 1358
Db 1116 -----GGGKK 1120
QY 1359 LRSTVOSTETGLAVERNMTROASRESTDGSNYSSEGNLTFPGVRLASDSQFDFL 1418
Db 1121 LRSTVOSTETGLAVERNMTROASRESTDGSNYSSEGNLTFPGVRLASDSQFDFL 1190
QY 1419 DGLGPAQLVGRQTLATPAMGDIVGMMDKKQLEVEIIRARGLVYKGSKTLPAIPYKVY 1478
Db 1181 DGLGPAQLVGRQTLATPAMGDIVGMMDKKQLEVEIIRARGLVYKGSKTLPAIPYKVY 1240
QY 1479 LLDNGVCIAKKTKVARKTLEPLXQULLSFEEBPQGRVLOIYWGDIYGRMDHKSFGVQAQ 1538
Db 1241 LLDNGVCIAKKTKVARKTLEPLXQULLSFEEBPQGRVLOIYWGDIYGRMDHKSFGVQAQ 1300
QY 1539 ILDELELENNVIGWFKLFPSSSLVPTSAFLTRRRASQSSLESSTGYSYSR 1590
Db 1301 ILDELELENNVIGWFKLFPSSSLVPTSAFLTRRRASQSSLESSTGYSYSR 1352

```

```

RESULT 8
Q9JIR9 PRELIMINARY; PRT; 1368 AA.
AC Q9JIR9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RIM2-3A.
GN RIM2-3A.
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP
SEQUENCE FROM N.A.
RX MEDLINE=20347919; PubMed=10748113;
RA Wang Y., Sugita S., Sudhof T.C.;
RT "The RIM/MIM Family of Neuronal C2 Domain Proteins. Interactions with
RT Rab3 and a new class of Src homology 3 domain proteins."
RL J. Biol. Chem. 275:20033-20044(2000).
EMBL: AF199325; AAF81647.1; -.
DR InterPro: IPR000008; -.
DR InterPro: IPR000306; -.
DR InterPro: IPR000345; -.
DR InterPro: IPR001478; -.
DR Pfam: PF00595; PDZ; 1.
DR ProSITE: PS50004; C2 DOMAIN_2; 2.
DR ProSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART: SM00239; C2; 1.
SQ SEQUENCE 1368 AA; 155559 MW; CE08E16F0B383A11 CRC64;

```

```

Query Match 78.5%; Score 6504; DB 11; Length 1368;
Best Local Similarity 78.8%; Pred. No. 0;
Matches 1283; Conservative 18; Mismatches 29; Indels 296; Gaps 6;
QY 1 MSAPLGPGRGAPRPAASQPPQPMPLSHLTEERKILIAVMDROKKEEKEQSVLKI 60
Db 1 MSAPLGPGRGAPRPAASQPPQPMPLSHLTEERKILIAVMDROKKEEKEQSVLKI 59

```

```

QY 61 KEENAOPTOMPFSGITELVNNVLOPOOKOPNEKEPOTKLHOOFENYKBOVKMGESQ 120
Db 60 -----KLHOOFENYKBOVKMGESQ 80
QY 121 QOEOBKDAPTCGICHTKTPKADCGCHNCYCOTKFCARCGGRVSLSNKVMWVCNLCRKQ 180
Db 81 QOEOBKDAPTCGICHTKTPKADCGCHNCYCOTKFCARCGGRVSLSNKVMWVCNLCRKQ 140
QY 181 QEILTKSAMWYNSGNTLQOPDQKVPRLRNEAEAOEKKAKLHEQPOFGPAGDLSVPA 240
Db 141 QEILTKSAMWYNSGNTLQOPDQKALRGLRSEAEAOEKKAKLHEQPOFGPAGDLSVPA 200
QY 241 VEKGAHILTRDITLKNKSGVHQIADMPSPDKRSPSVSDQNRREOGESEEDYQOYV 300
Db 201 VERGAHILTRDITLKNKSGMKGHQLASDMPSPDKRSPSVSDQNRREOGESEEDYQOYV 260
QY 301 PSDGTPSPSDYADRSQREPOFYEEPHLNYRDSNRGRHRSKEYIVDEDEVESRDEY 360
Db 261 PSDGTPSPSDYADRSQREPOFYEEPHLNYRDSNRGRHRSKEYIVDEDEVESRDEY 320
QY 361 ERORREEEYQARYRSDPNLARYPVKPOPYEEOKRIAEVSRABERHRSOVSLANALED 420
Db 321 ERORREEEYQARYRSDPNLARYPVKPOPYEEOKRIAEVSRABERHRSOVSLANALED 380
QY 421 SRISLLRMDRPSRORSVSEBRAMENQASYSMERTREAOQSSYPORTSNHSPPTPRSP 480
Db 381 SRISLLRMDRPSRORSVSEBRAMENQASYSMERTREAOQSSYPORTSNHSPPTPRSP 440
QY 481 IPLRDPMRADSLKROHHLDPSSAVALKTKREKMETLRLNDLSLSDQSESVPRPPPHK 540
Db 441 IPLRDPMRADSLKROHHLDPSSAVALKTKREKMETLRLNDLSLSDQSESVPRPPPHK 500
QY 541 SKKGKMKOVSLSSSEELASTPEYTSODVLESESVSKGDSQKCKRTSSEGVLSDS 600
Db 501 SKKGKMKOVSLSSSEELASTPEYTSODVLESESVSKGDSQKCKRTSSEGVLSDS 560
QY 601 NTRSEOKKRMYYGHSLEDEWSEPOIKDSGVDTCSSTLNEHSHSDKHPTWQPSK 660
Db 561 NTRSEOKKRMYYGHSLEDEWSEPOIKDSGVDTCSSTLNEHSHSDKHPTWQPSK 620
QY 661 DGDRLIGRILLNKLKDSGVPRDPSGAMLGKLVGKMTESGRICAFTTKVKKGSLADTVG 720
Db 574 DGDRLIGRILLNKLKDSGVPRDPSGAMLGKLVGKMTESGRICAFTTKVKKGSLADTVG 633
QY 721 HLRPDDEVLENNGRLLQGTFFEEVYNIILSKPEQVELVSRPISGIDPRIPDTHAOLE 780
Db 634 HLRPDDEVLENNGRLLQGTFFEEVYNIILSKPEQVELVSRPISGIDPRIPDTHAOLE 693
QY 781 SSSSFEESQKMDRPSISVTPSPMSPGMLRDVPOFLSGOLS-----IKLMF 824
Db 694 SSSSFEESQKMDRPSISVTPSPMSPGMLRDVPOFLSGOLS-----IKLMF 753
QY 825 DKGHQLITVTLIGANDLPREDGPRPNPVYKIFLPRDSQKMKRTKTKLEPKMNQ 884
Db 754 DKGHQLITVTLIGANDLPREDGPRPNPVYKIFLPRDSQKMKRTKTKLEPKMNQ 813
QY 885 FTYSVYHRRERFERMELTLMDOARVRESESEFLGELLILETALDDEPRMYLQTHDV 944
Db 814 FTYSVYHRRERFERMELTLMDOARVRESESEFLGELLILETALDDEPRMYLQTHDV 873
QY 945 SSLPRLRPSPYLPRRLOHGESPTRLORSKRISDSEYSDVDCEDGCVGVSVYRNGRDLQ 1004
Db 874 SSLPRLRPSPYLPRRLOHGESPTRLORSKRISDSEYSDVDCEDGCVGVSVYRNGRDLQ 933
QY 1005 SSTLSVPEQVWSSNHCSPGSPHRYDVIGRTKRSWSPSPAPPORNVGEGHGRATGHN 1064
Db 934 SSTLSVPEQVWSSNHCSPGSPHRYDVIGRTKRSWSPSPAPPORNVGEGHGRATGHN 993
QY 1065 ISRMRRHVMMDHYSRDR-----RCEADROPIHHSRSTEQ 1102
Db 994 ISRMRRHVMMDHYSRDR-----RCEADROPIHHSRSTEQ 1053

```

```

QY 1103 RPLERTTTRSSRSERPDNTLMKSMPSLMTGRSAPPSPALSRSHPRTGSVQTSPTSPT 1162
Db 1054 RPLERTTTRSSRSERPDNTLMKSMPSLMTGRSAPPSPALSRSHPRTGSVQTSPTSPT 1113
QY 1163 GRRGROLPOLPRKGTLEBERAMDIIEERNRQMKLNKYKOVASDPRLEDDYHSKYSGMDPH 1222
Db 1114 GRRGROLPOLPRKGTLEBERAMDIIEERNRQMKLNKYKOVASDPRLEDDYHSKYSGMDPH 1131
QY 1223 RGADIVSTKSSDSVDVSAVSRSTSSASRPSSTSYMSVQSERPRGNKIKSVFTSKMQNRQ 1282
Db 1132 ----- 1131
QY 1283 MGVSGKNLTSTISIGDMCSLEKNDSQSDPTAVGALGTSGKRRSSIGAKWVAIVLSRK 1342
Db 1132 ----- 1131
QY 1343 SRASASOLSOTBEGGKLRSTVQSTETGLAVENRNMWTRQASRSTDSGSMNSYSSEGLI 1402
Db 1132 -----GGKKLRSTVQSTETGLAVENRNMWTRQASRSTDSGSMNSYSSEGLI 1180
QY 1403 PPGVRLASDQSFDFLDGLGPAQLVGRQTLATPAMGDIQVGMMDKKGQLEVEIIRANGLY 1462
Db 1181 PPGVRLASDQSFDFLDGLGPAQLVGRQTLATPAMGDIQVGMMDKKGQLEVEIIRANGLY 1240
QY 1463 YKPSKTLPPAPYKVVYLLDNGVCIARKKTKVARKTLEPLYQOLSFEEESPQGRVLOIIVW 1532
Db 1241 YKPSKTLPPAPYKVVYLLDNGVCIARKKTKVARKTLEPLYQOLSFEEESPQGRVLOIIVW 1300
QY 1532 GDYGRMDKHSFNGVQAOILDELELSMNVTGFKLFPSSSLVDPLSAPLTRASOSSLESS 1582
Db 1301 GDYGRMDKHSFNGVQAOILDELELSMNVTGFKLFPSSSLVDPLSAPLTRASOSSLESS 1360
QY 1583 TGPYSYRS 1590
Db 1361 TGPYSYRS 1368

```

RESULT 9
 09JTR8 PRELIMINARY; PRT: 1292 AA.
 AC 09JTR8: 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
 DE RIM2-4B.
 GN RIM2-4B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20347919; PubMed-10748113;
 RA Wang Y., Sugita S., Sudhof T.C.;
 RT "The RIM/NIM Family of Neuronal C2 Domain Proteins. Interactions with
 R1 Rab3 and a new class of Src homology 3 domain proteins.";
 RL J. Biol. Chem. 275:20033-20044(2000).
 DR EMBL: AF199328; AAF81650.1; -;
 DR InterPro: IPR000008; -;
 DR InterPro: IPR000306; -;
 DR InterPro: IPR000345; -;
 DR InterPro: IPR001478; -;
 DR Pfam: PF00168; C2; 2;
 DR Pfam: PF00595; PDZ; 1;
 DR PROSITE: PS50004; C2_DOMAIN_2; 2;
 DR PROSITE: PS500190; CYTOCHROME_C; UNKNOWN_1;
 DR SMART: SM00239; C2; 1;
 SQ SEQUENCE 1292 AA; 146933 MW; E15B2C2B25452094 CRC64;

Query Match 75.0%; Score 6212; DB 11; Length 1292;
 Best Local Similarity 77.1%; Pred. No. 0;
 Matches 1226; Conservative 21; Mismatches 45; Indels 298; Gaps 5;

```

OY 1 MSAPLPGRRPAPTPAASOPPOPEMDLSHLEERKILAVNDROKKEEKEOSVLI 60
DB 1 MSAPLPGRRPAPTPAASOPPOPEMDLSHLEERKILAVNDROKKEEKEOSVLI 59
OY 61 KEHKAOPTQWPFSGITTELNVNVLQPOKOPNEKEBOTKLHOOFEMKKEOVKKMGESQ 120
DB 60 -----KLHOOFEMKKEOVKKMGESQ 80
OY 121 QOOEQGADPTGICHTKTRADCGCHNCSTCYTKFCARCGGRVLSRKNVAVCNLCRKQ 180
DB 81 QOOEQGADPTGICHTKTRADCGCHNCSTCYTKFCARCGGRVLSRKNVAVCNLCRKQ 140
OY 181 QELLTSGAMFYSGSNTLQOPDOKVPGRJLNEBAPQEKAKLHEOPFOGAGLSTVA 240
DB 141 QELLTSGAMFYSGSNTLQOPDOKVPGRJLNEBAPQEKAKLHEOPFOGAGLSTVA 200
OY 241 VERGRAHGLTRQDTIKNGSGVKHQIASDMPDRKRSPSVSRDQNRKRYEQSEEREDYSQYV 300
DB 201 VERGRAHGLTRQDTIKNGSGVKHQIASDMPDRKRSPSVSRDQNRKRYEQSEEREDYSQYV 260
OY 301 PEDGTPRBPSPDYADRRSQEPQFYEEPHLNTYDQNSNRGRHSHKEYIVDDEDVSRDEY 360
DB 261 PEDGTPRBPSPDYADRRSQEPQFYEEPHLNTYDQNSNRGRHSHKEYIVDDEDVSRDEY 320
OY 361 ERQRRREERYQARYSDPNLARVPYKPOPYEEQMRTHAEVSRARHERRSHDVLAAELED 420
DB 321 ERQRRREERYQARYSDPNLARVPYKPOPYEEQMRTHAEVSRARHERRSHDVLAAELED 380
OY 421 SRISLLRMDRPSQORSVSERRAAMENORSYSMERTREAOGOSYPQRTSNHSPPTPRSP 480
DB 381 SRISLLRMDRPSQORSVSERRAAMENORSYSMERTREAOGOSYPQRTSNHSPPTPRSP 440
OY 481 IPLDRPMRRAADSLRQOHLDPSAVRKTREKMETLRNDLSLSDQESVSRPPRRPHK 540
DB 441 IPLDRPMRRAADSLRQOHLDPSAVRKTREKMETLRNDLSLSDQESVSRPPRRPHK 500
OY 541 SKKGCMROVSLSSSEELASTPEYTSQCDVLELESSEVSSEKGSOKGKRKTSSEQVLSDS 600
DB 501 SKKGCMROVSLSSSEELASTPEYTSQCDVLELESSEVSSEKGSOKGKRKTSSEQVLSDS 545
OY 601 NRSERQKRMYTGGHSLSEEDLEWSEPOIKDQGVDTCSSTTLNEHSHDKHRYTWQPSK 660
DB 546 -----YSWLEHSHDKHRYTWQPSK 573
OY 661 DDDRILGRILLNKRLKDSVPRDSGAMLGLKVVGGKMTESGRLCAFTTKVKKSLADTVG 720
DB 574 DDDRILGRILLNKRLKDSVPRDSGAMLGLKVVGGKMTESGRLCAFTTKVKKSLADTVG 633
OY 721 HLRPGDEVLEWNGRLQGAFFEEVYNILLESKPEPQVELVSRPIGDIPIRIPDSTHAOLE 780
DB 634 HLRPGDEVLEWNGRLQGAFFEEVYNILLESKPEPQVELVSRPIGDIPIRIPDSTHAOLE 693
OY 781 SSSSSSESGKMDRPSITSVSPMSPGMLRDVPOFLSGQLSIKLMDVKYGHOLIIVTLGAKD 840
DB 694 SSSSSSESGKMDRPSITSVSPMSPGMLRDVPOFLSGQLSIKLMDVKYGHOLIIVTLGAKD 753
OY 841 LPSRDEGRBNPNVKTLYFLPDRSDKNKRRKTKYKTLLEPKMNQFFITSVPVRRERREML 900
DB 754 LPSRDEGRBNPNVKTLYFLPDRSDKNKRRKTKYKTLLEPKMNQFFITSVPVRRERREML 813
OY 901 ETLMDQAVRESESEFLGILLLEETALDDDEPHWYKLOTHDVSSLPLPPSPYLPARQ 960
DB 814 ETLMDQAVRESESEFLGILLLEETALDDDEPHWYKLOTHDVSSLPLPPSPYLPARQ 873
OY 961 LHGESPTRLQSRKSLSDSEVSDYDCEDEGVVYSDYRHNCRDLOSSLVPEQVWSSNHC 1020
DB 874 LHGESPTRLQSRKSLSDSEVSDYDCEDEGVVYSDYRHNCRDLOSSLVPEQVWSSNHC 933
OY 1021 SPGSGHRAVDVIGRTSWSSAPPPORANVQGHGRTATGHYNTISRMDRHRVMDHYSS 1080
DB 934 SPGSGHRAVDVIGRTSWSSAPPPORANVQGHGRTATGHYNTISRMDRHRVMDHYSS 993

```

```

OY 1081 DRRDCEAADRQPYHRGRSTEQRLBERTTTRSSRSERPDTNLRSMPSLMTGSRAPSP 1140
DB 994 ERDSHFLTLPRSRHRQTSSEHHRD----- 1017
OY 1141 ALSRSHPRGTSVQTSPPSTPGTGRGRQOLPOLPKGLTERSAMDIEERNQMKLKYQOV 1200
DB 1018 --GRSHPRGTSVQTSPPSTPGTGRGRQOLPOLPKGLTERSAMDIEERNQMKLKYQOV 1055
OY 1201 AGSDPRLQDYHSKYRSGMDPHKAGADTVSTKSSDSDVSDVSAVSTSSASFSTSYMSV 1260
DB 1056 ----- 1055
OY 1261 QSEPRGRNRIQVFTSKMQRQMGVSGKNLTKSTISGDMCSLEKNDGSDTAVALGT 1320
DB 1056 ----- 1055
OY 1321 SGKRRSSIGAKVAIVGLSRKSRASQLSQTEGGKKLRSTVORSTETGLAVERNMT 1380
DB 1056 -----GGKKLRSTVORSTETGLAVERNMT 1082
OY 1381 RQASRESTGSMNSYSSEGNLIFPGVRLASDSQPSDFLDGLGPAQVGRQTLATPAMDI 1440
DB 1083 RQASRESTGSMNSYSSEGNLIFPGVRLASDSQPSDFLDGLGPAQVGRQTLATPAMDI 1142
OY 1441 QVGMADKKGQLEVEIIRARGLVYKPGSKTLPAPYKYVLLDNGVCIAKKTKVARKTLEP 1500
DB 1143 QVGMADKKGQLEVEIIRARGLVYKPGSKTLPAPYKYVLLDNGVCIAKKTKVARKTLEP 1202
OY 1501 LYQQLLSFEESPOGKVLQIIVWGDYGRMDKSPGVQAQIILDELELSNMVIGWFKLEPPS 1560
DB 1203 LYQQLLSFEESPOGKVLQIIVWGDYGRMDKSPGVQAQIILDELELSNMVIGWFKLEPPS 1262
OY 1561 SLVDPSTAPLTRASQSSLESSGSPYSRS 1590
DB 1263 SLVDPSTAPLTRASQSSLESSGSPYSRS 1292

```

```

RESULT 10
090026 PRELIMINARY; PRT; 1188 AA.
AC 090026;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE KIA0751 PROTEIN.
GN KIA0751.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=9087487; PubMed=9872452;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT For large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
DR EMBL; AB018294; BA034471.1; -.
DR HSSP; P31016; 1BE9.
DR InterPro; IPR000008; -.
DR InterPro; IPR001478; -.
DR Pfam; PF00168; C2; 2.
DR Pfam; PF00595; PD2; 1.
DR PROSITE; PS00004; C2 DOMAIN_2; 2.
DR SMART; SM00239; C2; 1.
SQ SEQUENCE 1188 AA; 135249 MW; DD65DB32288055AF CXC64;

```

Query Match 68.1%; Score 5638; DB 4; Length 1188;
Best Local Similarity 83.3%; Pred. No. 0;


```

Db 177 EDKVVMMVNCNCRKQOELLTKSGAMFGSGP---QOPSQD---GTLSDTATGAGSEVPRE 230
QY 219 KKAKLHPOFOGAPGLSPVAVEKGAHGLTROTITKNGS-----GKXH 263
Db 231 KKAKLOERSNSQ---TPLSTAAY-----SSQDTATPGAPLHRNKGAEPSQALPQEQ 279
QY 264 QIA---SDMPSRKRRSPSVSRDQNRREYEQSEEBEDYSQY--PSDQTPMRSPSDYADR 317
Db 280 KQASRSRSEEPREKKAHGLS-EDNGKGQKSEKKRVKSVQGEI-----ADERERK 333
QY 318 SOEPFOYEEPCGLNTRDSNRKGRHSREXYIVDEVEDESREYERORREERYOARYSDP 377
Db 334 ERRETRLEKRSQSDYSD---RPEKRDNGRAVD-----QKQKEEYOTRYISDP 381
QY 378 NLARYPVKPOPYEEOKRIHAEVARHERHSHVSLANALEDSRSLSLMRDPSRQSV 437
Db 382 NLARYPVKAPPEOQKRMARVSRARHERHSHVAPLPHNE-----AAAAAP 427
QY 438 SERRAAMENORSTSMERTREAOQSSYPORTSNHSPPTPRSPPIPLDRPDRARADSLRQ 497
Db 428 ABAATGKRAPATARVSPSPSPRARAATAAPPTTEHGPPPPRPAPAGRAPEPEPRVPEPLRQ 487
QY 498 HHLDPSSAY--RKTREKMETMLRNDLSLSDQSESVRPPPPRHKSCKGKMROVSLSS 555
Db 488 GLDLDGSAVLLKAKRKEKESMLRNDLSLSDQSESVRPPPHRPRKRGKRRQMSVSS 547
QY 556 EEBELASTPEYTSQCDVLESESVSEKGSQKGRKRTSEQVLSDSNTRSEBQKKRYAYG 615
Db 548 EEEGVSTPEYTSCEDEVELESESVSEKGDLD-----YY-- 579
QY 616 HSLIEDLPMSEPOINDSGVDTCSSTTLNEHSHSKHPVTWQPSKDGRLIGRIILNKL 675
Db 580 -----WLDP-----ATWHSRETSPISSHPTWQPSKEGRLIGRIILNKT 620
QY 676 KDGVPKDSGAMGLKVYGGKMTESGRLCAFTTKYKGLADVTGHLRPGDVLWNGRL 735
Db 621 ---TMPKESGALLGLKVYGGKMTDLGRGAFITTKYKGLADVGHGLRAGDVLWNGRP 677
QY 736 LOGATFEVYVITLESREPEYELVSRPIDRIPRDPSTHAOLESSSSSEFSQKMDRPS 795
Db 678 LPGAFTNEEVNITLESSEPEVEITIVSRPIDRIPRESSHPLESSSSSEFSQKMDRPS 737
QY 796 ISVTSFMSPGMLRDVPOFLSGOLSIKLMFDKVGHQLYITLIGAADLPSREGCRPNPYK 855
Db 738 ISVTSFMSPGMLKDAQVLPQOLSVKLMYDKVGHQLYITLIGAADLPRVYGRPNPYK 797
QY 856 IYFLPDSDNKRRKTKVTKLEPKNQFTLYSPVHRERERMLETIMQDARVREPES 915
Db 798 MYFLPDSDSKRRKTKVTKLEPKNQFTLYSHVHRDRFRERMLEITWMDQPRVODEES 857
QY 916 EFLGELLIELETAALLDDEPHMYKLOTHDVSLPRPSPIPLRQOLGSESTRLOSKR 975
Db 858 EFLGELLIELETAALLDDEPHMYKLOTHDVSLPRPSPIPLRQOLGSESTRLOSKR 917
QY 976 ISDSSEVDYCEGEGVYV--DYRHNGRDLQSSSTLSVPEQVMSSNHCSPGSPHRYVIG 1033
Db 918 ISSDLSIDYEVDGIGVVPVYRASARESKATTLTVPEQOTTHHSRVSYHRKGDQ 977
QY 1034 RTRSMSPSAPPPORANVEOGH--RGTRATGHYNTISRM---DHRVMDHYSQD----- 1081
Db 978 RPRSRPLPNV-PLQSRSLDEIHPTRRSRSPTRHDSRSPADHRSRHVESQYSSPDSLELM 1036
QY 1082 -----RDRDCEA----- 1088
Db 1037 LPRAKRGSAESLHMTSELQPSLDRARASATNCLRPDTSLHSFERHRHSKSEKSCIOQO 1096
QY 1089 -----AD-----KQYHRSRSTQRPDLLETRTTRSSSERPDTNLMSPSLMT 1132
Db 1097 SRKGTASDADRVLPCPLSLRGYATPRATQDPVYRGKYPTRSSSEHSSVRLTCSMHMLAP 1156
QY 1133 GRSAPSP-ALSRSHPRTGVSQTSSTSPGTRGGRGOLPOLPK-GTLERSAMDLIEERN 1190
Db 1157 GGSAPSPULLLRTHRQSPQSPPADTSFGSRGRGOLPOLPOVYRSGSIEQASLVEEERT 1216

```

```

QY 1191 QM--KLMKYOV--AGSDPRLLEDYHSKYRSGMDPHRGADIVYTKSSDSDYSAVSRT 1246
Db 1217 QMKVYHFRKQOTSSGSSQOELDHEQSKYNTIHKDQYRSCDNASAKSSSDYSDVSAISRA 1276
QY 1247 SSARFESTSYMSYOSEPPRRNRKRTISVFTSKQNRQKMGVSGKNLTSTISGDMCSEKN 1306
Db 1277 SSTRLSTSTMSQOSERPRG--RISFTPMQGRMGTSGRAIITKSTVSGETIYTLERN 1334
QY 1307 DGSQSDTAVALGATSGKRRSSIGAKMVAIVGLSRKSRASASOLQOTGGGKLTSTVORS 1366
Db 1335 DGSQSDTAVALGATGAGKRRSSLSAKVAVIY--SRBRSTISQLQOTSGHKKLSTIORS 1392
QY 1367 TEGLAIVEMNMWMTROKRSRSTSDGSMNSYSEGNLIFPGVRLASQSPDLGLGPAOL 1426
Db 1393 TEGMAEMKRR-MKQPSRSTSDGINSYSEGNLIFPGVAVGPDQSPDLGLGPAOL 1451
QY 1427 VGRQTLATPAMGDIQVGMMDKQOLEVEIIRARGLVYKPSKTLTPARYVYLLDNVCI 1486
Db 1452 VGRQTLATPAMGDIQGMEDKQOLEVEIRARSLTQKPSKSTPARYVYLLDNVCI 1511
QY 1487 AKKTKVARKTLEPLVQQLSFEESPOGRVLOITWGDYGRMDHKSFMGYAQLLDEL 1546
Db 1512 AKKTKVARKTLDPLVQQLSVPDESPOGKLVQYIWMGDYGRMDHKGFMGYAQLLELD 1571
QY 1547 SNMVIQMFLLPSSLVDPFSAPLTRRASQSSLESSTGSPYSRS 1590
Db 1572 SNMVIQMYKLFPSSLVDPFSAPLTRRASQSSLESSTGSPYSRS 1615

RESULT 12
ID 035168 PRELIMINARY; PRT: 1553 AA.
AC 035168;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DR 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE RAB3 EFFECTOR.
GN RAB3
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN.
RX MEDLINE=97394473; PubMed=9252191;
RA Wang Y., Okamoto M., Schmitz F., Hofmann K., Sudhof T.C.;
RT "Rim is a putative Rab3 effector in regulating synaptic-vesicle
RT fusion."
RL Nature 388:593-598(1997).
DR EMBL: AF007836; AAB66703.1; -.
DR HSSP: P21707; IRSY.
DR InterPro: IPR000008; -.
DR InterPro: IPR000194; -.
DR InterPro: IPR000306; -.
DR InterPro: IPR001360; -.
DR InterPro: IPR001478; -.
DR InterPro: IPR003315; -.
DR Pfam: PF00168; C2; 2.
DR Pfam: PF00595; PD2; 1.
DR Pfam: PF02318; RPH3A_effector; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE: PS00004; C2_DOMAIN_2; 2.
DR PROSITE: PS00572; GYCOSTYL_HYDROL_FL_1; UNKNOWN_1.
DR SMART: SM00239; C2; 1.
SQ SEQUENCE 1553 AA: 172964 MW: D664397E65D03BE CRC64;

```

Query Match 54.2%; Score 4489.5; DB 11; Length 1553;
 Best Local Similarity 57.2%; Pred. No. 8.8e-288;
 Matches 952; Conservative 209; Mismatches 319; Indels 183; Gaps 37;


```

QY 1 MSAPLPGRRPAAPPAASQPPPOPEMDLSHLTEERKILLAVMDROKKEEKEQSVLKI 60
Db 1 MSSAVGPGRRPPT-----VPPMQELDLSHLTEERNITIMAVMDROKKEEKEEAMLK 56
QY 61 KEENAO-----TQFPFSGITELVNNVLOPOOKOPNEKEPOT--KLHOOFEMIKKE 110
Db 57 VVRDMAKPAACKTPPNASQPHQPPPLNFRVCVPRKPSSEEGGPEMDMLHOOFESYKE 116
QY 111 QVKKMGESOOQO--BOKGAPTCGICHTKTFADGCGHNCSCYOTKFCARCGRVSLSNK 169
Db 117 QVKTIQGEARRKQGHKKDDAPTCGICHTKTFADGCGHLCSTCYTKFCARCGRVSLSNN 176
QY 170 ---VWVCNLCRKQOEILTKSGAFYNGSNTLQOPDOKVPRGLRNE-----EAPQE 218
Db 177 EDKVMVWVCNLCRKQOEILTKSGAFYNGSNTLQOPDOKVPRGLRNE-----EAPQE 230
QY 219 KKAKEHPOPOFGAGDLSVPAVEKGRANGLTRQDTIKNGS-----GVKH 263
Db 221 KKAKEHPOPOFGAGDLSVPAVEKGRANGLTRQDTIKNGS-----GVKH 279
QY 264 QIA-----SDMPDRKRSPSVRDONRREYOSFEREDYSQV--PSDGTMRSPSDYADR 317
Db 280 KOASRSRSEPPERRKAPOLS--EONGKGQKSERKRPKSVYOPGEGT-----ADEREK 333
QY 318 SQREPOFYEEPHLNYRDSNRGRHRSKEYIVDEDEVESKDEXERQREBEYOARYSDP 377
Db 334 ERRETRLEKGRSQYSD--RPEKRDNGRAVED-----QKQKEEYQTRYRSDP 381
QY 378 NLARKPVAPPOPEEOMRIHAESRAHERRHSDVSLANAELEDSRISLMDRPSRORSV 437
Db 382 NLARKPVAPPOPEEOMRIHAESRAHERRHSDVSLANAELEDSRISLMDRPSRORSV 447
QY 438 SERRAMENORSYMERTRERAGOSSTPQRTSNHSPPTPRSPILDRPMRADSRKQ 497
Db 428 AEATAGKAPATARKVSPSPSPARAAAAPPTHEGPPPPRPAAGPAPPPRPPRPELRKQ 487
QY 498 HHLDPSSAV--RRTREKMETMLRNDLSLSDQSESVRPPRPHKSKGKMRQVSLSS 555
Db 488 GRLDGSAVLLAKKAREKESMLRNDLSLSDQSESVRPPRPHKSKGKMRQVSLSS 547
QY 556 EEELASTPEYITSCDDVLESESVSEKDGOKKRTSEQCVLSDSNTSRQKKRMYYGG 615
Db 548 EEEGVSTPEYITSCDDVLESESVSEKDGOKKRTSEQCVLSDSNTSRQKKRMYYGG 579
QY 616 HSLIEDLMSEPOKIDSGVDTCSSTTLNEHSHSDKPHYTWQPSKODRLIGRLIKRL 675
Db 580 -----WLDP-----ATWHSRETSPISHPTWQPSKEDRLIGRYILKRT 620
QY 676 KDGVPDRSGAMLGLKVVGGKMTESGRLCAFTTKVKKGSLADTVGHLRPGDEVLEWNGRL 735
Db 621 ---TWPKESGALLGLKVVGGKMTDLGRGALFTTKVKKGSLADTVGHLRAGDEVLEWNGRP 677
QY 736 LQGAFFEEVYNTILLESKPEPOVELYVSRIGDIPRTDSTHAOLESSSSSEFESQKMDRPS 795
Db 678 LPGAINEEVYNTILLESKPEPOVEIIVSRIGDIPRIPESSHPLSESSSSSEFESQKMDRPS 737
QY 796 ISVTPSPMGMLRDVPOFLSGOLSTIKLWDFKGHOLIVITLGAOKLPSRDGPRNRYVK 855
Db 738 ISVTPSPMGMLRDVPOFLSGOLSTIKLWDFKGHOLIVITLGAOKLPSRDGPRNRYVK 797
QY 856 IYPLPDRSDKMKRTKTYVKKLEPKWNTQFTIYSPVHRRERREMLLETTLMDQARVEES 915
Db 798 MYPLPDRSDKMKRTKTYVKKLEPKWNTQFTIYSPVHRRERREMLLETTLMDQARVEES 857
QY 916 EFLGEIILELTALLDDPRHWKLOTHTVSSLPPLRPSPLPRKOLHGESPTRLQSKR 975
Db 858 EFLGEIILELTALLDDPRHWKLOTHTHSSSLPLPQSPFMPRRRHJGESSSKRLQSKR 917
QY 976 ISDSEVSDYDCEDGVGVYS--DYRHNGRDLQSTLSVPEQVMSNCSPPSGPHRVYIG 1033
Db 918 ISDSDISDYEVDGIGVYPRVGYRASARESKATITLVPEQQTTHHRSKRVSHRGDQG 977
QY 1034 RTRSWSPSAPPQORNEVOGH--RGTFRATGHYNTISRM---DRHRVMDHXYSSDRDRCEA 1088

```

```

Db 978 RPRSLPUNV--PIQRSIDEIHPTRRSRSPTRHNDASSPADHRSHRVESQXSESPDSILM 1036
QY 1089 ADROPTNRS-----RTEORPLLERTTTSRSSERPDTNL---MRSPSLMT----- 1132
Db 1037 LPRARRGSAESLHMHTSELQPSLDRARSASTNCLRPTSLASPERENHRSKSERCSIOKO 1096
QY 1133 GRSAPPSPALSRSHPRGTSVQTSPPSSTPGTGRGRROLPOLPK--GTLERAMDIEERNQ 1191
Db 1097 SKKGTASDA--DHTHROGSTQSPADTSFGSRGRROLPOVPVYVSGSLTEQASLYVEETRQ 1155
QY 1192 M--KLNKYQV--AGSDPRLIEDYHNSKYRSGWDPHRRGADTVSTRSSDSDDVSDVASRST 1247
Db 1156 MKVYHRRFQOTGGSSQGLDHEQYSKYNIHKDOYRSCDNASAKSSDSDVSDVASIRAS 1215
QY 1248 SASRFSSTSYMSQVSRPRGNKRKISVFTSMQNRQKQVSGSKNLTKSTISGDMCSLEKD 1307
Db 1216 STSRLSSTSFMSQSERPRG--RISSETPRMOGRRMGTSGRAIKTSYSGEITLERND 1273
QY 1308 GSQSDTAVALGTSGKKRRSSIGAKMVAIYGLSRKRSASQLSQTEGGKKIRSTVQRT 1367
Db 1274 GSQSDTAVALGTSGKKRRSSIGAKMVAIYGLSRKRSASQLSQTEGGKKIRSTVQRT 1331
QY 1368 ETGLAVEMRMNTROASRESTDGSNNSYSSEGNLIPGVRLASDSQSFDELGLGPAQV 1427
Db 1332 ETGMAEMRK--MVQPSRSTDSQINSYSSEGNLIPGVRLASDSQSFDELGLGPAQV 1390
QY 1428 GQOTLATPAMGDIQVGMAMKQGLEVEIIRAGLVYKPGSKTLPAPVYKYLLDNGVIA 1487
Db 1391 GQOTLATPAMGDIQVGMAMKQGLEVEIIRAGLVYKPGSKTLPAPVYKYLLDNGVIA 1450
QY 1488 KKTIVARKTEPLVQOLSFEESSPOGRVLOLITVWGDYGMHDKSFNGVQOILDELELS 1547
Db 1451 KKTIVARKTEPLVQOLSFEESSPOGRVLOLITVWGDYGMHDKSFNGVQOILDELELS 1510
QY 1548 NNVIGWFKLEPPSSLVDPSTAPLTRRASQSLSSSTGSPSYSRS 1590
Db 1511 NNVIGWFKLEPPSSLVDPSTAPLTRRASQSLSSSTGSPSYSRS 1553
RESULT 13
O15048 ID O15048 PRELIMINARY; PRT; 1053 AA.
AC O15048;
AD O15048;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE KIA0340 (FRAGMENT).
GN KIA0340.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N. A.
RC TISSUE=BRIN;
RX MEDLINE=97349984; PubMed=9205841;
RA Tanaka T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
DR EMBL; AB002338; BAA20798.1; -.
DR HSSP; P21707; IRSY.
DR InterPro; IPR000008; -.
DR InterPro; IPR000306; -.
DR InterPro; IPR001478; -.
DR InterPro; IPR001565; -.
DR InterPro; IPR003315; -.
DR Pfam; PF00168; C2, 1.
DR Pfam; PF00595; PDZ, 1.
DR Pfam; PF02318; RPH3A_effector; 1.

```

DR PRINTS: PR00399; SYNAPOTAGMN.
DR POSITE: PS50004; C2_DOMAIN_2; 1.
DR SMART: SM00239; C2; 1.
FT NON_TER 1
SQ SEQUENCE 1053 AA; 117818 MW; D31DD27271F5E981 CRC64;

Query Match 35.4%; Score 2936.5; DB 4; Length 1053;
Best Local Similarity 54.1%; Pred. No. 2e-185;
Matches 614; Conservative 143; Mismatches 232; Indels 145; Gaps 24;

```
QY 1 MSAPLPGRRAPPAASOPPPMPDLSTLTERKTIILAVMDRÖKKEEGSVLKI 60
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  26 MSSAVGGRGPPPT---VPPMOELPDLSTLTEREERNITAMVMDRÖKKEEGEEMLC 81
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||

QY 61 KEHKADPTQWFPFSGITELVNNVLQPOÖKÖPNEKEPQTKHOOPMYEÖVKKMGEEG 120
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  82 VVRDMARPA-----CKTPRAENQPHQPSRLHQEFSEYKÖVAKTIGEAR 128

QY 121 QÖQ-EÖKGDAFTGCIHKTKFKADGCGHNCYCÖTKFCARCGGVSLSRNK---VMMVCN 175
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  129 RYÖGEHMDAPTCGICHTKFRADGCHLCYÖTKFCARCGGVSLSRNNDKVVMMVCN 188

QY 176 LCRKQÖELTTSGAMFYSGSNTLQÖPQÖVPRGLRNE-----EAPÖKKAHLHÖQÖ 228
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  189 LCRKQÖELTTSGAMFYSGSNTLQÖPQÖVPRGLRNE-----EAPÖKKAHLHÖQÖ 228

QY 229 FÖ-----GAPGDLSPVAVEKGRAGHL--TPÖDITKNSGVKHÖASMPDRKRSYV 279
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  243 SÖPLSTAAASSÖDAPPAAPPDRSKGAEPÖQALGPEÖKQASSRSRSPPRKKTPEGL 302

QY 280 SRDÖNRREYÖSEEREDYSÖYVPSDÖTMRPSDÖYADRSÖRÖPEYÖEFGHLNYRDSNR 339
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  303 SEÖGKGALKEERR---VPKTSQÖVEGAVEERERER-----RESRL 344

QY 340 GHRSKETIYDÖDEVESDÖ-----YÖRÖREYÖYÖARÖPNLARYVÖKÖYÖEÖQR 394
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  345 EKGRSÖDÖ---PDPRPEKDEKKADEKÖRKEDYÖTRYSÖPNLARYVÖKÖPEÖQÖQR 401

QY 395 IHAVSARHERRRHSÖVSLAÖELEDÖSILRMDRPSÖRÖSVSEERAMENÖRÖSMER 454
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  402 MHAVSARHERRRHSÖVSLAÖELEDÖSILRMDRPSÖRÖSVSEERAMENÖRÖSMER 454

QY 455 TREÖGÖSSYÖÖRTSNÖSPPTÖRSPÖPLDÖRMRADÖSLRÖKHÖLDPÖSÖVÖRÖKREK 514
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  458 TÖRÖRÖRÖGÖ---KÖLÖTNSÖPÖRÖGÖVÖPÖAÖEÖLÖKÖEÖLÖRÖSLDÖPÖSÖV 507

QY 515 EYMLRNDÖSLDÖSÖESVÖRPPÖRPHKÖSKÖGÖKÖMÖVÖSLSSÖEÖELASTÖPEYÖSCDÖVELE 574
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  508 -LÖMRNDÖSLDÖSÖESVÖRPPÖRPHKÖSKÖGÖKÖMÖVÖSLSSÖEÖGVSTÖPEYÖSCDÖVELE 566

QY 575 SESYSEKGDÖSGÖKGRKÖTSEÖGVÖLSDÖSNÖRSEÖRÖKÖRMÖYÖGHSÖLEDEÖLMSÖQÖTDSÖV 634
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  567 SESYSEKGDÖL-----TY-----WLDP----- 583

QY 635 DÖCSÖTLNÖEÖSHÖDKÖPÖTÖWÖPSÖKDÖRLÖIGÖRLÖKÖLÖKDÖSVÖRDSÖGÖMLÖGLÖVÖG 694
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  584 ---ÖTÖMÖSRÖTÖPSÖLSSÖPÖTÖWÖPSÖKÖEGÖDLÖIGÖRYÖLÖKÖT---ÖTÖPKÖSGÖLLÖGLÖKÖVÖG 636

QY 695 GÖKÖTESÖGLÖCAFÖTKÖVÖKÖGÖSLÖADÖYÖGÖLÖRÖPGÖDEVÖLÖMÖGÖRLÖÖGÖATÖFEEVÖNYÖLÖLESÖKEP 754
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  637 GÖKÖMÖDLÖRLÖGÖAFÖTKÖVÖKÖGÖSLÖADÖYÖGÖLÖRÖPGÖDEVÖLÖMÖGÖRLÖÖGÖATÖFEEVÖNYÖLÖLESÖKEP 696

QY 755 PÖYÖELVÖYÖRÖPÖGÖLÖPÖRÖPÖTÖHÖAÖEÖSSÖSÖFÖEÖKÖMDRÖPÖSTÖVÖSPÖSGÖLÖRÖDÖVÖQÖL 814
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  697 PÖYÖELVÖYÖRÖPÖGÖLÖPÖRÖPÖTÖHÖAÖEÖSSÖSÖFÖEÖKÖMDRÖPÖSTÖVÖSPÖSGÖLÖRÖDÖVÖQÖL 756

QY 815 SÖGÖLSÖIKÖLÖPÖDÖVÖGÖHÖLÖVÖTÖLÖGÖAKÖDÖLPSÖREDÖGRÖPÖNYÖVÖKÖYÖLÖPÖDÖSÖDKÖKÖRÖKÖTÖYÖK 874
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  757 PÖGÖLÖSVÖLÖMÖDÖVÖGÖHÖLÖVÖTÖLÖGÖAKÖDÖLPSÖREDÖGRÖPÖNYÖVÖKÖYÖLÖPÖDÖSÖDKÖKÖRÖKÖTÖYÖK 816

QY 875 KÖTÖLÖPÖKÖNÖTÖFYÖSVÖHÖRÖFERÖRÖMÖLÖTÖTÖLÖDÖQÖARÖVEÖSEÖFÖLÖEÖLÖTÖLÖDÖP 934
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
```

```
DB 817 KÖLÖPÖKÖNÖTÖFYÖSVÖHÖRÖFERÖRÖMÖLÖTÖTÖLÖDÖQÖARÖVEÖSEÖFÖLÖEÖLÖTÖLÖDÖP 876

QY 935 HÖWKÖLÖTHÖDÖVÖSLÖPLÖPÖSPÖYLÖPÖRÖLÖGÖEÖSPTÖRLÖQÖSKÖRÖLÖSÖEÖVDYÖDÖCEÖDVÖGVYS 994
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  877 HÖWKÖLÖTHÖDÖVÖSLÖPLÖPÖSPÖPÖRÖRÖHÖGÖESSÖKÖLÖRÖSÖRÖLÖSÖDÖSDYÖEÖDÖGÖVÖP 936

QY 995 -DÖYÖHÖNGÖRDÖLÖSÖSTLÖYÖPEÖVÖMÖSSÖNHÖCSÖPSÖPÖHÖVÖDÖVÖIÖTRÖRÖSMÖSÖAPPÖQÖHÖVÖGÖ 1052
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  937 PÖVÖGRÖSSÖARÖESÖKÖSTLÖVÖYÖEÖQÖRTÖHÖHÖRSÖVÖPÖHÖGÖDÖQÖKÖPÖRÖLÖPNV--PÖLÖRÖSLÖDEI 995

QY 1053 H--RÖGTÖRÖTÖGHÖYÖNTÖISÖRÖ--DÖHÖRÖVMDÖHÖYÖSDÖRDÖCEÖADÖRÖPYÖHÖRÖSTÖE 1101
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  996 HÖPÖRÖRSÖSPÖRÖHÖDÖASÖRPÖYÖDÖRÖTÖRDÖVÖDÖYÖLÖSÖDÖSÖELÖMLÖPÖRÖ--RÖGRÖSÖAE 1047

DB 996 HÖPÖRÖRSÖSPÖRÖHÖDÖASÖRPÖYÖDÖRÖTÖRDÖVÖDÖYÖLÖSÖDÖSÖELÖMLÖPÖRÖ--RÖGRÖSÖAE 1047
```

RESULT 14
ID 09HBA5 PRELIMINARY; PRT; 740 AA.

AC 09HBA5;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE RAB3 INTERACTING PROTEIN VARIANT 2 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Coppola T., Maguin-Luthi S., Perret-Menoud V., Gattesco S.,
RA Regazzi R.;
RT "Molecular characterization of six different isoforms of RIM (Rab3
interacting molecule) from human brain.";
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF263306; AAG23163.1; -.
FT NON_TER 1
SQ SEQUENCE 740 AA; 83352 MW; 09CFB918F8FE8063 CRC64;

Query Match 29.3%; Score 2424.5; DB 4; Length 740;
Best Local Similarity 64.9%; Pred. No. 7.9e-152;
Matches 493; Conservative 105; Mismatches 125; Indels 37; Gaps 12;

```
QY 848 RPRNPYKIFLÖPÖRSÖKNÖRÖKÖTÖVÖKÖTÖLÖPÖKÖNÖTÖFYÖSVÖHÖRÖFERÖRÖMÖLÖTÖTÖLÖDÖP 907
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  1 RPRNPYKIFLÖPÖRSÖKNÖRÖKÖTÖVÖKÖTÖLÖPÖKÖNÖTÖFYÖSVÖHÖRÖFERÖRÖMÖLÖTÖTÖLÖDÖP 907

QY 908 ARVÖEÖSEÖFÖLÖEÖLÖTÖLÖDÖDEÖHÖWÖYÖKÖLÖQÖHÖDVÖSSLÖPÖRÖSPÖYLÖPÖRÖLÖGÖEÖSPT 967
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  61 PÖVÖEÖEÖSEÖFÖLÖEÖLÖTÖLÖDÖDEÖHÖWÖYÖKÖLÖQÖHÖDVÖSSLÖPÖRÖSPÖYLÖPÖRÖLÖGÖEÖSPT 120

QY 968 RÖLÖÖRÖKRÖIDÖSEYÖSDYÖDÖCEÖDVÖGVYS--DÖYÖHÖNGÖRDÖLÖSÖSTLÖYÖPEÖVÖMÖSSÖNHÖCSÖPSÖGS 1025
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  121 KÖLÖÖRÖSÖRÖIDÖSDYÖDÖCEÖDVÖDÖGÖVÖGVYÖPÖVÖGÖYÖSÖARÖESÖKÖSTLÖVÖYÖEÖQÖRTÖHÖHÖRSÖSVS 180

QY 1026 PÖHÖRÖVÖDÖGÖTÖRÖSMÖSPÖAPÖPÖQÖHÖVÖEÖGÖH--RÖGTÖRÖTÖGHÖYÖNTÖISÖRÖ--DÖHÖRÖVMDÖHÖYÖS 1080
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  181 PÖHÖGÖNDÖGÖKÖPÖRÖSLÖPNV--PÖLÖRÖSLÖDEIÖHÖPÖTÖRÖRSÖSPÖRÖHÖDÖASÖSPÖYÖDÖHÖRÖDÖVÖDÖYÖLÖS 239

QY 1081 DÖRDÖCEÖADÖRÖPYÖHÖRÖS-----RÖSTÖQÖRÖPÖLÖERTÖTÖRÖSSÖRÖEÖPÖDÖTÖNLÖMRÖSPÖLÖMTÖGRS 1135
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  240 EÖDÖSÖELÖMLÖPÖRÖKÖGRÖSÖAEÖCHÖHTÖSÖLÖPÖFLÖDRÖARÖSÖATÖNÖLÖRÖDÖTÖL----- 287

QY 1136 APPÖSPÖLÖSRÖSHÖPÖTÖGÖSÖVÖTÖSPÖSPTÖGÖRÖGÖLÖPÖLÖPÖK--GÖTÖLÖRÖSÖMDÖIÖEÖRÖNRÖKÖL 1194
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  288 --HÖSPÖERÖMÖHÖRÖSÖPÖYÖSPÖADÖTÖSÖFSÖSRÖGÖQÖLÖPÖYÖVÖRÖGÖSÖIÖQÖASÖLÖVÖEÖEÖRÖQÖKM 345

QY 1195 --NÖKYÖQV--ÖGÖSPÖRÖLÖEÖYÖHÖSKÖYÖRÖSGÖWPÖHÖRÖGÖADÖYÖTÖKÖSSÖDÖVÖSÖDÖVÖSÖAVÖSÖTÖSSAS 1250
  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||  ||:|||||
  346 KÖYÖHÖFÖKÖTÖGÖSSÖQÖEÖLÖDÖRÖYÖSKÖYÖNÖHÖKÖDÖYÖRÖSCÖDÖVÖSÖASÖSDÖVÖSÖDÖVÖSÖAVÖSÖTÖSSAS 405

QY 1251 RÖFSÖTÖSYÖMÖSÖVÖSÖERÖPÖRÖNRÖISÖVÖTÖSKÖMÖNÖRÖMÖGÖVÖSGÖKÖNLÖKÖSTÖSÖISÖGÖMÖSÖLÖEÖKÖDÖSÖQ 1310
```

```

Db 406 RLSTSFMSGESEPRG--RISSEFTPKMOGRMGTSGRSLMKSTSVSGEMYTLEHNDGSQ 463
1311 SPTAVGALCTSGKKRRSSIGAKMVAIVGLSKRSASQSLQTEGGKKLASTVORSTETG 1370
464 SPTAVGTGAGGKKRRSSISAKVVAIV--SRSRSTQSLQTESGHHKLLKSTIORSTETG 521
1371 LAVERNMWTROASRSTSGSMNSYSEGNLIPPGVRLASDSQSPDLGIGPAQLVGRQ 1430
522 MAEMRK -MYRQPSRESTDGSINSYSEGNLIPPGVRLGADSPDFDLGIGPAQLVGRQ 580
1431 TLATPAMGDIQVGMDDKQGLEVEIIRARGLYVKGSKTLPAPYKYLNDNGVCIARKK 1490
581 TLATPAMGDIQIGMEDKKQGLEVEIRARSLTQKPSKSTPAPYKYLLENGACIAKKK 640
1491 TKVAKTLEPLVQQLSFESPOGRVLOIIVMGDYGMDHKSFMGVAQIILDELELSMV 1550
641 TRIAKRTLDPLVQOSTLVFPESSPOGKVLQVIVMGDYGMDHKCFMGVAQIILDELELSMV 700
1551 IGMFKLFPPSSLVDPSTAPUTRRASQSSLESSTGSPYSRS 1590
701 IGMFKLFPPSSLVDPPLTLPLTRASQSSLESSTGSPCIRS 740

```

```

RESULT 15
ID Q9HBA6 PRELIMINARY: PRT: 766 AA.
AC Q9HBA6:
BT 01-MAR-2001 (TREMBLrel. 16, Created)
BT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
BT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE RA3 INTERACTING PROTEIN VARIANT 1 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_maxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RA Coppola T., Maguin-Luthi S., Perret-Menoud V., Gattesco S.,
Regazzzi R.;
" Molecular characterization of six different isoforms of RIM (Rab3
interacting molecule) from human brain."
Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF263305; AAG23162.1; -.
FT NON TER 1
SQ SEQUENCE 766 AA; 86537 MW; CF492313B221C3D0 CRC64;

```

Query Match 29.1%; Score 2410.5; DB 4; Length 766;
 Best Local Similarity 63.0%; Pred. No. 7e-151;
 Matches 495; Conservative 105; Mismatches 123; Indels 63; Gaps 14;

```

Cy 848 RPRNRYKLYFLPDRSDKKRTKTVKTLPEKWNQTFIYSPVHRREFRRLMLETIMDQ 907
1 RPRNRYKMYFLPDRSDKSKRRKTKVTKILEPKWNQTFYVSHVRDFRRLMLETIMDQ 60
908 ARVREESFLEGLIETLALDDPRHWKTLQTHVSSLPRLRPSYLLPRRLHGESPT 967
61 PRVQEESEFLEGLIETLALDDPRHWKTLQTHDESSLPLPQSPFMPRRNTHIGSSS 120
968 RRLQSKRISDESVSDCEDGVGVVS--DYRNGRDLQSTLSVPEQVMSNNHCSPGS 1025
121 KKLQSRQISDSDISDIYEVDDGIGVVPVGYRKSARSKSTTLTLVPEQRTTHHRSVS 180
1026 PIRVDVIGTRSMSPAPPQRNVQGH--RGTATGHTNTISM--DRHRVMDHYSS 1080
181 PIRGNDQGRPRSLRPV-PLQSLDEIHPTRRSRSPTRNHDASRPVDHRTROVDSQYLS 239
1081 DRD-----RDCA-----ADROPYHRS-----TEORPLLERT 1109
240 EODSELMLPRAKRGSAECGLTTRHLVRYHTLPPKMPLOSSSHWNYSLELOPFLDRA 299

```

```

Qy 1110 TTRSRERPDNTLMRSMPSLMTGRASPAPSPALSRSHPTGVSOTSPSPGTGRGRL 1169
300 RASPTNCLRPDTSL-----HSPEERMRHROSPTQSPADTSFSSRRGRL 345
1170 POLPK-GTLERSAMDIEERNQML--NKYKOV--AGSDPRLQDHYHSKYRSGMDPRG 1224
346 PQVPVRSQIEQASLVYEERTROMKMHVRFKQTTGSGSSQELDRQYSKYHNKQYRS 405
1225 ADYVTRKSSDSVDVSAVSRSTSSAFSSTSYMSVQSERPRGRKRTSVFTSKNQNRMG 1284
406 CUNVASKSSDSVDVSAISRSTSSASRLSTSEMSSEPRG--RLSSFTPKMQGRMG 463
1285 VSGKNLTKSTISGDMCSLEKNDGSDQTAVALGTGKKRRSSIGAKMVAIVGLSKRSR 1344
464 TSGRSTMKSTSVSGEMYTLEHNDGSQSDTAVGVYAGAGKKRRSSLAKVVAIV--SRSR 521
1345 SASQSLQTEGGKKLRSTVORSTETGLAVEMRMWMTROASRESTDGSMSNSYSEGNLIF 1404
522 STQSLQTESGHHKLLKSTIORSTETGMAEMRK -MYRQPSRESTDGSINSYSEGNLIF 580
1405 GYRLASDSQSPDFDLGIGPAQLVGRQTLATPAMGDIQVGMDDKQGLEVEIIRARGLYV 1464
581 GYRLGADSPDFDLGIGPAQLVGRQTLATPAMGDIQIGMEDKKQGLEVEIRARSLTQK 640
1465 PGSKTLPAPYKYLNDNGVCIARKKTVAKTLEPLVQQLSFESPOGRVLOIIVMGD 1524
641 PGSKTLPAPYKYLLENGACIAKKTRIKRKTLDPLVQOSTLVFPESSPOGKVLQVIVMGD 700
1525 YGRMDHKSFMGVAQIILDELELSNMVIGWFKLPPSSLVDPSTAPUTRRASQSSLESSTG 1584
701 YGRMDHKCFMGVAQIILDELELSMVIGWFKLPPSSLVDPPLTLPLTRASQSSLESSTG 760
Qy 1585 PYSYRS 1590
761 PPCIRS 766

```

Search completed: November 21, 2001, 16:09:59
 Job time: 265 sec

Fri Nov 23 10:08:44 2001

us-09-617-099b-2.rge

FD/PO 7/14 (2000)
Rev. 2 + Ref (JBC) Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2001, 15:57:54 ; Search time 8528.31 Seconds
(without alignments)
9032.199 Million cell updates/sec

Title: US-09-617-099b-2
Perfect score: 4980
Sequence: 1 gcttcctaggggtgttcgg.....acattgtgcctagcaaaag 4980

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_com: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_p11: *
13: gb_p12: *
14: gb_p13: *
15: gb_p14: *
16: gb_da1: *
17: gb_da2: *
18: gb_fun: *
19: gb_htgo_hum: *
20: gb_htgo_inv: *
21: gb_htgo_rdg: *
22: gb_htg_hum1: *
23: gb_htg_hum2: *
24: gb_htg_hum3: *
25: gb_htg_hum4: *
26: gb_htg_hum5: *
27: gb_htg_hum6: *
28: gb_htg_hum7: *
29: gb_htg_hum8: *
30: gb_htg_inv1: *
31: gb_htg_inv2: *
32: gb_htg_other: *
33: gb_htg_rdg: *
34: gb_hum1: *
35: gb_hum2: *
36: gb_hum3: *
37: gb_hum4: *
38: gb_hum5: *
39: gb_hum6: *
40: gb_hum7: *
41: gb_in: *
42: gb_com: *
43: gb_or: *

44: gb_ov: *
45: gb_pat: *
46: gb_ph: *
47: gb_p1: *
48: gb_p10: *
49: gb_p11: *
50: gb_p12: *
51: gb_p13: *
52: gb_p14: *
53: gb_da1: *
54: gb_da2: *
55: gb_fun: *
56: gb_htgo_hum: *
57: gb_htgo_inv: *
58: gb_htgo_rdg: *
59: gb_htg_hum1: *
60: gb_htg_hum2: *
61: gb_htg_hum3: *
62: gb_htg_hum4: *
63: gb_htg_hum5: *
64: gb_htg_hum6: *
65: gb_htg_hum7: *
66: gb_htg_hum8: *
67: gb_htg_inv1: *
68: gb_htg_inv2: *
69: gb_htg_other: *
70: gb_htg_rdg: *
71: gb_hum1: *
72: gb_hum2: *
73: gb_hum3: *
74: gb_hum4: *
75: gb_hum5: *
76: gb_hum6: *
77: gb_hum7: *
78: gb_hum8: *
79: gb_in1: *
80: gb_in2: *
81: gb_in3: *
82: gb_com: *
83: gb_ov: *
84: gb_pat1: *
85: gb_pat2: *
86: gb_ph: *
87: gb_p11: *
88: gb_p12: *
89: gb_p13: *
90: gb_p14: *
91: gb_da1: *
92: gb_da2: *
93: gb_fun: *
94: gb_htgo_hum: *
95: gb_htgo_inv: *
96: gb_htgo_rdg: *
97: gb_htg_hum1: *
98: gb_htg_hum2: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4608.4	92.5	4800	AB021131	AB021131 Mus muscu
2	3886	78.0	5550	AF199331	AF199331 Rattus no
3	3834	77.0	5592	AF199330	AF199330 Rattus no
4	3776	75.8	5640	AF199322	AF199322 Rattus no
5	3197.6	64.2	5172	AF199324	AF199324 Rattus no
6	2936.4	59.0	4965	AF199329	AF199329 Rattus no
7	2860.4	57.4	5031	AF199323	AF199323 Rattus no
8	2860.4	57.4	5031	AF199326	AF199326 Rattus no

C	45	135.2	2.7	67106	81	AC090666	Homo sapi
C	44	136	2.7	187350	78	AL512289	Homo sapi
C	43	136	2.7	187457	80	AL357652	Homo sapi
C	42	136	2.7	129040	92	HS1046C13	Homo sapi
C	41	171.8	3.4	166384	75	AC074113	Homo sapi
C	40	176.2	3.5	864	95	AF0002081	Homo sapi
C	39	175.2	3.5	175443	61	AC0099710	Homo sapi
C	38	181.8	3.7	559	94	G57511	Homo sapi
C	37	185.4	3.7	1163	94	MM0310531	Homo sapi
C	36	187.2	3.8	176976	62	AC0512213	Homo sapi
C	35	193.8	3.9	245	58	G05813	Homo sapi
C	34	201.4	4.0	283078	82	AL5590011	human STS W
C	33	211.4	4.2	167766	83	AC0028849	human STS W
C	32	271.2	5.4	154485	73	AC0686934	human STS W
C	31	302	6.1	185460	64	AP0016832	human STS W
C	30	314	6.3	167769	83	AF263308	human STS W
C	29	377	7.6	1791	89	AF263309	human STS W
C	28	377	7.6	1551	89	AF263310	human STS W
C	27	377	7.6	1605	94	DB0704	human STS W
C	26	379.6	7.6	1565	94	AF199334	human STS W
C	25	398.2	8.0	1424	91	BC003103	human STS W
C	24	398.2	8.0	7239	91	BC003103	human STS W
C	23	683.6	13.7	2103	89	AF263307	human STS W
C	22	715	14.4	6395	85	AB002338	human STS W
C	21	741.6	14.9	162489	91	AP001572	human STS W
C	20	741.6	14.9	139451	69	AC025836	human STS W
C	19	761.6	15.3	2301	89	AF263305	human STS W
C	18	782.4	15.7	1663	94	AF199335	human STS W
C	17	796	16.0	2323	89	AF263306	human STS W
C	16	940.2	18.9	1877	88	AF007156	human STS W
C	15	1038.8	20.9	5841	94	AF199333	human STS W
C	14	1204	24.2	5655	94	AF007836	human STS W
C	13	2364.8	45.7	3854	85	AB018294	human STS W
C	12	2631.2	52.8	4851	94	AF199332	human STS W
C	11	2802.4	56.3	5079	94	AF199335	human STS W
C	10	2860.4	57.4	5031	94	AF199327	human STS W
C	9	2860.4	57.4	5031	94	AF199327	human STS W

ALIGNMENTS

```

gene
CDS

/cell_line="MINE"
32..4624
/ene="RIM2"
32..4624
/ene="RIM2"
/codon_start=1
/product="Rim2"
/protein_id="BAB18975.1"
/db_xref="GI:11611475"
/translation="MSAPGRGRAPPPAASQPPPEPMDLSLTHEEKRTILAVI
DQKKEEKEQSVLKIKEKHQAQPTMPFFSITELVNVNLOQKOPKEPPORTIL
QOFEMKIEQVKKKKEESQOOQKGDGAPGCHTKRAGGQKSCQKRCRCC
GAVSLRNSVMVMVNCILKCRQOEILTKSGAFNFSQNSNLTQDQKVPVGLRNEA
PKAKLHEOPOFOGAPGDLSTVAPKEKRGAGLITROTINKGSVFKHIDMSDEKRR
PSVSRDQNRRTYQSEEREDYSOYVSDSGMPSPSDVADRSQKQIPQYEIEGHILN
DSNRGHRNRKSTIVDDDEVSDEDEYEQKREEEYOARYRSPNINARVPKQPEE
MTHAEVSRARERHRHSDVSLANAELEDSRISLEIMDRPSRVSERRAMENORS
SMERREMOGQSSYSPORTSNHSPPTPRPSIYELDRDMKRDVLSLKHHLDPSSAVR
TKREMEFTMLRNDSSDOSSESVRPPRPHKSKGKGMDVSLSSKEELASPEEY
SCDVELSESSEYSEKGDQKGRKSGEGLVSDNSTREROKKRRATYGHSLIEDLE
SEPOLKDSGVDTCSSTTLNEHSHSDKRPVYTMQPSKSDHRLIGRLILMKRLKDSVP
DSGAMLGKLVYGGKRTESGRLCARITTKKASLAVTGLHPRGVEVLWENRLLQGA
PEEYNILTESPEPQVELVAVSRIGDLPRIPIDSHQALESSESFESEKMDRPSYK
TSPMSGMLRDVPQPLSGQLSIKLMFDVGVQLVYTLIGARDLPSREDDGRPNRYK
YELPDRDNKRRTYKTKLEPKNQOTFISPVHIREFERMEITLMDQARYEESE
SEFLFEILIELELAILDEDPHYMKLOTQVDSLPLPRSPPLTROLGSEPTROLQ
SKRISSEVSDYDCDEGQVGVSDYRHNDRLOISLSTPDEVMASSNHCSPGSPHRYLQ
VIGRISSESPAPPRORVNEBOGHNRGTRATGHTNITSKDRHRKMDHYSDDRSHP
TGSVOTSPSTPTGTORGRQLQLPQPKOTLRSSAMDIEERNRQKLNKXYKQVADSDP
LEQDYHAKTRISGMDPHRGADTVSTKSSDSDVDSVAYSRTSSASRSTSYMSQSEER
PRGNKRTISVFTSKMQNRQMGVSGKNLKTSISGDGMSEKNDSSQDSTAVGALGTSG
KRRKRSIGAKKAVYLVGLSRKSRASQLSQTEGGGKKLSTYORSTETGLAVEMKMTG
RQASRETSQDSMNSYSEGNLIPGVRLASDSQPSDFLDGPAQIIVGROTLATPMAG
DIQMGMDKQGLEVEIIRAGLVKRGSKTLPAPVYKVLVNDGVCYIAKKTYVARK
TLEPLYOOLISFEESPOGRVLIQITVMYDGMKDRKSMGVAQIILDELSNMYIGMEF
KLPPSSVPTLAPILTRASQSSLESSTGVSYS"

```

RESULT 1
 AB021131
 LOCUS 4800 bp mRNA ROD 20-DEC-2000
 DEFINITION Mus musculus RIM2 mRNA, complete cds.
 VERSION AB021131
 KEYWORDS AB021131.1 GI:11611474
 SOURCE RIM2.
 ORGANISM Mus musculus cell_line:M16 cDNA to mRNA.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sates)
 Ozaki,N., Shibaas,T., Kashima,Y., Miki,T., Takahashi,K., Ueno,H.,
 Sunaga,Y., Yano,H., Matsura,Y., Iwanaga,T., Takai,Y. and Seino,S.
 CAMP-GEFT is a direct target of CAMP in regulated exocytosis
 Nat. Cell Biol. 2 (11), 805-811 (2000)
 11056535
 2 (bases 1 to 4800)
 Ozaki,N. and Seino,S.
 Direct Submission
 Submitted (12-DEC-1998) to the DDBJ/EMBL/Genbank databases. Nobuaki
 Ozaki, Nagoya University School of Medicine, First Department of
 Internal Medicine, Nagoya, Nagoya 466-8550, Japan
 (E-mail:n-ozaki@med.nagoya-u.ac.jp, Tel:81-52-744-2142,
 Fax:81-52-744-2157)
 Sequence updated (22-Jun-1999)
 Sequence updated (02-Feb-2000).
 Location/Qualifiers
 1..4800
 /organism="Mus musculus"
 /db_xref="taxon:10090"

	Query Match	92.5%;	Score 4608.4;	DB 94;	Length 4800;
	Best Local Similarity	96.4%;	Pred. No. 0;		
	Matches 47/9;	Conservative	0; Mismatches	1; Indels	180; Gaps
QY	1 gcttccttaaggtggttgcgtcccgccaacatgttcgctcgtcggccccgggccc	60			
Db	1 GCCTCCATGGGCGGTGGCTCCGCCAACAATGTCGGTCCGCTCGGGCCCCGGGCGC	60			
QY	61 cccgcgtcccaaccgccggcgctctloaacctctccgaagcccgatgpcgagactcag	120			
Db	61 CCCGGCTCCACACC CGGGCGCTCTCAACTCTCCGAGCCGAGATGCCGACTCTAG	120			
QY	121 ccacctcaagaaaggagaagaatacatctctgtctctatgatatcttgaagaaga	180			
Db	121 CCACCTCAAGGAAGAGAGAGAATAATCTCTGGCTGTATGATGTCGTCAAAAPAAA	180			
QY	181 agagagaagaagagcagtcocgtgtcctaagatcaaagaacaacaagcacacgcaca	240			
Db	181 AGAGGAAGAAGGAGCGTGCTCGTCAATCAATAAAGAACACAAAGCACCAACCGACACA	240			
QY	241 gtgtgttcctccttta tgggatacactcgaactctgttaataagctttcgagcccccagcaaaa	300			
Db	241 GTGTGTTCCTTTA TGTGGGATTCCTGAATCTGTAATTAACGTTCTGCAGCCCCAGCAA	300			
QY	301 acaaaccaatyaagaagagccccagacaagaagtgcacacaacnaatttgaaatglaaaya	360			
Db	301 ACAACCCAATAGAAGAGGCCCCAGACAAACCTGCACCAACAATTGTAATGTAAGA	360			
QY	361 gcaagtcagaagaatttgagagagagatcgcaacagcagagaagaagcagaaggtgatgcccc	420			
Db	361 GCAAGTCAAGAAGATGGGAGAGGATTCGACACACACGCAAGCAGAGAGGTGATGCCCC	420			
QY	421 gacctgtgcatctgtccacaagaacaaatttgcagatgtagtcgagcataattttctcta	480			

|||||
Db 421 GACCTGTGGCATCTGCACAGACAAATTTTGAGATGATGCGGCGCATATTTGTTCTTA 480
Qy 481 ttgcacaacaaagttctgtgtcgcgatgtgaggtcgaagtgtcttcttaagctcaaaagtg 540
Db 481 ttgccaacaaacgaattctgtgtcgcgatgtgaggtcgaagtgtcttcttaagctcaaaagtg 540
Qy 541 tatgtgggtgtgtaattttgtgcgaaacacaaagaatctctacbaatccagagatg 600
Db 541 TATGTGGGTGTATTTTGTGCCGAAAACACAGAAATCTCTACTTAATCAGAGATG 600
Qy 601 gtttataatagttgggtctaacacactgcagcaacctgatalcaaaaggttccctcgaaggt 660
Db 601 GTTTATATATAGTGGGTCTAACACACTGCAGCAACTGATCAAAAGGTTCTCTGAGGGCT 660
Qy 661 tcgaaatgaagagacccctcgaagagaagaagaacaaactacagagacagcccaagttcca 720
Db 661 TCGAAATGAGAGAGCCCTCCAGGAGAGAAAGCAAACTACAGAGAGCCCAAGTTCCA 720
Qy 721 aggaagcccaagttgacttaacagtaacctgtgaagaaggccgaagctcatgggtcac 780
Db 721 AGGAGCCCAAGTGTACTTATCTAGTACTCTGCAGTTGAGAAAGGCCGAGCTCATGGGCTCAC 780
Qy 781 aagaaggaatactatlaaanaatggaatcgaagagtgaaagcaccagatltgcagtgacatgcc 840
Db 781 AAGACAGGATACTATTTAAATGAGATCAGAGAGTGAAGACACAGATTGCCAGTGACATGCC 840
Qy 841 ttcaagaagaanaagcaagtccatcagttgtccagagatcacaatccgaagatacagaagaag 900
Db 841 TTCAACACGAAAAACCAAGCCATCATGATGTCAGGGATCAAAATCGAAGATCGAGCAAG 900
Qy 901 tgaagaagaagaagactacacacagtatgttccctcagatgtfacaatgtcacaagatctcc 960
Db 901 TGAGAAAGAGAGAGACTACTCTACAGTATGTTCTTTCAGATGGTACATGCCAGATCTCC 960
Qy 961 ttccgatatgtctgaatagaagatctcgaagcttgagccctcaattatagaagaacctgttca 1020
Db 961 TTCGATATGCTGTATGAGAGATCGATTCAGCGTGAAGCTCAATTTATGAGAAACCTGTGCA 1020
Qy 1021 tttaaattacagagatcttaacagaagaaggccaatgacatctccaagaagtatatgttga 1080
Db 1021 TTTAAATTTACAGGGATTTCTTACAGAGAGGCCCATGACATTTCCAAAGATATATTGTGGA 1080
Qy 1081 tgaagaagaatgtggagagaagaatgaatataagaagaacaaaggagaagaatcca 1140
Db 1081 TGATTAAGATGTGGAGAGAGAGATGAATATGAAGAACAAAGAGAGAGGAGATACCA 1140
Qy 1141 ggcagctacagaagtgatccaaatctggccccgttaccgttaagccacaacctcaga 1200
Db 1141 GGCAGCTACAGAAAGTGAATCTGGCCCGGTATCCCGTAAAGCCACAACCCCTACGA 1200
Qy 1201 agaacaatgcatccatccagctgtgaagtggtccagggcagaacatgtagagaagcaagtgta 1260
Db 1201 AGAACAAATGCCCATTCACAGCTGAGGCTGCCAGGGCACGACATGAGAGAGACACAGTGA 1260
Qy 1261 tgttctttggaagaagcgtgaactagaagaatccagaatcttctctgtctaagatgtgtag 1320
Db 1261 TGTTCCTTTGGCAAAAGCGTGAAGTGAAGATTCACAGATTTCTCTGCTAAGATGGATAG 1320
Qy 1321 accatcaaggcaagaatctgtatctgaacgttagagctgcaatggaanaacaaacgatgtta 1380
Db 1321 AACATCAAGGCAAAAGATCTGTATCTGAAGGTAGAGCTGCAATGGAANAACCAACGATGTGA 1380
Qy 1381 ttcaatggaagaactcgaagagcttcagggacaagaatcttcatccacaagaagacctcaaa 1440
Db 1381 TTCAATGGAAGAAGAACTCGAGAGGCTCAGGAGACAAAGTTCTTATCCAAAGAGACCTCAAA 1440
Qy 1441 tcatagttctccaccctccgagcgagacctatacgcgtgtgtagaccacaagatgaagcg 1500
Db 1441 TCATAGTCTCTCCACCCCTCGCGGAGCCCTATACGCTTGTATAGACCAACATGAGAGCG 1500
Qy 1501 cgctactccctcctcaggaacagacacacttagatcccaagctcgtcgtgtgagaaagaa 1560
|||||

Db 1501 CGCTACTCCTCTACGGAACAGACACCATTAGATCCAGCTCTGCTGTGAGAAAGAA 1560
Qy 1561 ggcgaagaaaaaatggaacacatgttlaaggaaatgtattcttgaagttcagaacagttccagtg 1620
Db 1561 GCGAGAAAAAATGGAACCATGATTTAAGAAATGATTTCTTTAGTTCCAGACAGTCCGAGTC 1620
Qy 1621 agtgaagccgcccccaagcctcataaatccaagaaggaggtlaaaatgtgccaggt 1680
Db 1621 AGTGAAGCCGCCCCCAACCAAGGCTCATTAATCCAGAAAGAGAGTAAATGCCCCAGT 1680
Qy 1681 ttcaatgaacagctcgaagagaagcttggaatccacaacctgaaglatacaagctgtatga 1740
Db 1681 TTCACGTGACACACTCGAGAGAGAGCTGGCATCCACACTGAGATATCAAGCTGTGATGA 1740
Qy 1741 tgttgagctcgaagaagcgaagtgtgagtgagaagaaggagacaagttcaaaaggaaaaaaa 1800
Db 1741 TGTGAGCTGTGAAGGCAAGATGTGAGTGAAGAAAGGAGCAAGTCAAAAGGAAAAAGAAA 1800
Qy 1801 aactagtagcagggagttttgtcagatctcaacccaagttcgaagacaaaagaagaag 1860
Db 1801 AACTAGTAGCAGGAGAGTTTGTGCGATTTCTAACACAGGCTCTGAGAGACAAAAGAAAG 1860
Qy 1861 gatgtactatggtggccactctttggaagaagatltggaatggtctgaagctcagatlaa 1920
Db 1861 GATGTACTATGTTGGCCACTCTTTGGAAGAGAGATTGGAATGCTGAGCCTCAGATTAA 1920
Qy 1921 ggaactctgggttagataactgtatgtatgacaacaccttaagaaggagacatgccaatga 1980
Db 1921 GGACTCTGGGGTAGATACCTGTATGTAGCACACACCTTAAAGAGAGAGATGAGCCATATGA 1980
Qy 1981 taagaacctgtgaaccttgcaagcatccaaagaatggaatcgcgcctaatgttgatatt 2040
Db 1981 TAAAGCACTGTGACCTGTGCGAGCATCCAAAGATGAGATGCGCTAATGTGCTATTTT 2040
Qy 2041 attaaataagcgtltaaagaatggaagtgtacactcgaagctcagaagcaatgtcgggct 2100
Db 2041 ATTAATAAGCGTTTAAAGATGGAGTGTACCTCGAGACTCAGAGACCAATGCTGGGCTT 2100
Qy 2101 aagaagttgagaaggaagaatgactgaatcgaatcgaactltgtgcatlttatccaaagt 2160
Db 2101 AAAGGTTGAGAGAGAAAGATGACTGAATCAGGTGCACTTGTGCAATTTATACCAAGT 2160
Qy 2161 aaaaaaagaagtttgaagctatctgtatagacatctttagacaaagttgaatcttga 2220
Db 2161 AAAAAAGAGAAATTTAGCTGATCTGTAGACATCTTTAGCCAGGTATGATGAGCTTTGGA 2220
Qy 2221 atggaatggagagctatltgcaagaagccacatlttgagaagtttacacatatctaga 2280
Db 2221 ATGGAATGGAGGCTATGTGCAAGGAGCCACATTTGAGGAAGTTTACAAATTAATTCTAGA 2280
Qy 2281 atccaaacctgaaccacaagaattgaagcttgttlttcaaggccaattggagatatcttag 2340
Db 2281 ATCCAAACCTGAACCAAGATGTGAGCTTGTGTTCAAGGCCAATGTGAATATCTCTAG 2340
Qy 2341 aatactgtatgaacagcagacacaactggaatccgaatcttcaagctatltgaatccaaa 2400
Db 2341 AATACTGTATGACAGCAATGACACAACTGGAATCGAATCGATTCAGCTCATTTGATCTCAA 2400
Qy 2401 aatgacagcttctctataltccggtlaccctcccaaccatgaagctccggtcgtcgaagatgt 2460
Db 2401 AATGAGCCGCTCTTATATATTCGTTACCTTCACCAATGAGTCTGCGCATGTGAGAGATGT 2460
Qy 2461 cccgcaagttcctatcttggaacagcttcaataaactatggtlttgaagaagttgttcaaca 2520
Db 2461 CCCGAGTTCTTATCTGAGACAGCTTCAATMAAACTATGTTCGCAAAAGTTGCTCACCA 2520
Qy 2521 gttgaatgtacaattttggagcaagaagatcccttccaggaagaatlggaagccaaag 2580
Db 2521 GTTGATAGTAAATTTTGGAGCAAAAGATCTCCCTCCAGGGAAGATGGAGAGCCTAAG 2580
Qy 2581 gaatccctatgttlaagaattacttctccagatagaagtgaataaaataagaagaagac 2640
Db 2581 GAATCCTTATGTTAAGATTTACTTCTCCAGATAGAAAGATTAATAAATAAGAGAAAC 2640
|||||

QY	2641	aaaaacagtc	aaagaaac	ctcttgga	accacaa	tggaac	cagactt	caattca	ttctctctgt	2700
Db	2641	AAAAAGTC	AAAGAAAA	CTTTGG	ACCCAA	TGAAAC	CAACACTT	CTATTTAT	TCTCTGT	2700
QY	2701	ccacccag	agaaatc	ctcgga	cgaa	tgctgga	aattac	ctcttgga	tcaagctagat	2760
Db	2701	CCACCCAG	AGAAATTC	CGTGAA	CGAATG	CTGGAA	AAATTA	CCCTTTGG	ATCAAGCTAAGT	2760
QY	2761	tcgagaag	agagacg	aaatctc	ttagag	agagatt	ttaattga	tttgaa	aaacagcttgc	2820
Db	2761	TGAGAGAAG	AGAGCGAA	TTCTTAG	GAGATTTTA	TAAATG	AAATGGA	AAACAGCTTTGCT		2820
QY	2821	agatgatg	agccgact	ggtata	agctgc	agacc	ccatgata	gtctc	ctcatgc	2880
Db	2821	AAGATGATG	AGCCGACTG	GATATAG	CTGACAC	CCATGATG	CTCTCAT	TGTCAC	TCC	2880
QY	2881	tcgcccct	ccccatc	ctcccc	ggag	gcgcctcc	atgga	agagcc	caacgcgcagct	2940
Db	2881	TGCCCCCT	CCCCATAT	CTGCCCCG	AGGACG	CTTCATG	GAGAGGCC	AAAGCCAG	CAGGCT	2940
QY	2941	gcaaaagt	tcgaag	aaataag	tacag	tgaagt	gtctga	ctacga	ctgcgagagatgc	3000
Db	2941	GCAAAAGT	CGAAGAA	TAAGTAG	CAGTAG	AGTGTC	GACTAG	CAGCTGC	AGAGATGCGT	3000
QY	3001	gggagatg	atgctc	gatata	tcgaca	caatg	gcgcgata	cttcca	agctccagctgtgcgt	3060
Db	3001	GGGAGATG	ATGCTCAT	TATTCG	ACAA	TGGCCCCG	ATCTTCAA	AGCTCCAG	CTGTGTGCGT	3060
QY	3061	gccagaca	agatcat	gtcatca	aatcat	gtgctga	ccataag	gggtctctc	atctcga	3120
Db	3061	GCCAGACA	AGATCATGT	CATCAAT	CATTGCT	GCATCAG	GGTCTCCT	ATCTGAG	ATGA	3120
QY	3121	tgcttata	gaagaa	gaacag	tcatgt	gtgc	ctagtg	ccccctc	ctctcca	3180
Db	3121	TGCTTATAG	GAAGAAC	GAAGCAT	GTGCTAG	TGCTAG	TGCCCCCT	CTCTCAA	AGATGTGA	3180
QY	3181	acaagg	gcaccc	gagga	gaacag	ctgtct	actg	gccttaca	acaacaa	3240
Db	3181	ACAGGGGC	ACCCGAGG	ACAGCTGT	CTACTG	GCCTAT	CAACACA	AAATTA	GGCAATG	3240
QY	3241	acaccc	gtgtca	tga	tga	cacact	ctctc	tga	taga	3300
Db	3241	ACACCGTG	ATGATG	ATGACCA	CTACTT	CTAGAT	TAGATG	ATGAC		3300
QY	3301	taga	gcacat	catca	gatcc	agata	cca	agaaca	acag	3360
Db	3301	TAGAGC	ACATC	ATCAGAT	CCATC	CTCTC	CTAGAG	CGAG	CCAC	3360
QY	3361	caaccc	gtccag	atcctct	gtga	agctt	ctgta	taaa	acctatga	3420
Db	3361	CAACCGTC	CGATCCT	CTGTGA	AGCTTCT	GTATGA	AAACCTAT	GATG	CTCTCAT	3420
QY	3421	aatgatg	aaaga	atctgc	ccctc	cttca	ctgc	ctta	tcga	3480
Db	3421	AAATGATG	AAAGATCT	GCCTCT	CTCACTG	CTCTCA	CTGCCT	CTCACTG	CTCACTG	3480
QY	3481	gtctgt	ccaga	caagcc	catca	agta	ctcc	ggga	aca	3540
Db	3481	GTCTGTCC	AGACAG	CCATCA	AGTACT	CCGGAC	AGACG	AGAGG	CCGAC	3540
QY	3541	acaagct	tcaca	caaag	gaacat	tggag	agaag	tgtct	atgata	3600
Db	3541	ACAGCTTC	ACACAA	GGAAATTT	GGAGAA	TGCTAT	GATATAG	AGAGAA	TGCA	3600
QY	3601	ccaatga	aaactta	aca	aaat	ta	caacag	gtag	ccgga	3660
Db	3601	CCAATGA	AACTTAACA	AAATTA	CAACAG	GTAGCCG	ATGAC	CCCGA	ATGAGCA	3660
QY	3661	ttaacat	tcga	agat	ctgc	tca	gat	tggat	tcaca	3720
Db	3661	TTAACATTC	GATATC	GTCTG	CAAGAT	TGGATTC	CAACAT	ATAGG	GGGAC	3720
QY	3720	ttacac	ttcga	agat	ctgc	tca	gat	tggat	tcaca	3780
Db	3720	TTACAC	TTTGA	AGATCT	GTCTG	CAAGAT	TGGATTC	CAACAT	ATAGG	3780
QY	3781	aatcagt	ctctta	catcca	aaatg	caaca	agata	tgg	gcgt	3840
Db	3781	AATCAGT	CTCTTAT	CATCA	AAATGCA	AAACAG	ACAGAT	TGGCG	GTGAG	3840
QY	3841	tcgttc	agac	ga	caac	atca	ctgtcc	gtca	aa	3900
Db	3841	TGTTTAC	GACGAC	CAACACT	ATCTG	CTGCTCA	ATCAG	AGCGCCG	AGAGCA	3900
QY	3901	gaccaaa	agccag	catcag	tggag	acata	gtctc	atctg	ga	3960
Db	3901	GACCAAA	AGCCAGCAT	CACTGAG	ACATGCT	CACTGAG	AAAGATG	AGCA	AGCAGCA	3960

Db	2880	TTTATAGGAAGACTTAGGCTCATGGCTGCGCTTATGTGTCCTTCCTCCCTTAAAGAAATGTGGGA	2939
QY	3181	acaagagcacccgaagggaacagctgctactctgacctaaacacacattagccgaatgga tag	3240
Db	2940	ACAGGGGCTCTGACGGGACACGTCGTACTGCGCATTTCAATACATTTAGCCGGATGGACAG	2999
QY	3241	acaaccgtlcalygaltygacctactctctcaatagagac-----	3280
Db	3000	ACACCGTCTCATGAGCAGACCATTTACTCTTCAGAGAGACACAGTCATTTCTCACTTACC	3059
QY	3281	-----agggattgtgaagc	3294
Db	3060	TCGCTCTCGACACAGGCAGACCAAGTAGCATCACCAAGGGATGAAAGGGATGTGTGAGC	3119
QY	3295	agcagaatagacagccataltacacagatccagatccaagaacaacgqccctccatagagc	3354
Db	3120	AGCAGATGACAGCCCATATCACAGATCCAGATACCAAGAAACACGGCCTCTCCTAGAGCG	3179
QY	3355	gaccacccaccgctccaga tccctctgaagctcc tga tacaacctca tgaagtcgalyc	3414
Db	3180	GACCAACACCCGCTCCAGATCCTCTGAACGTGCTGATACAAACCTCATGAGGTGATGCC	3239
QY	3415	ttcatatgaactggaagatctgcgccctccctcaactgccttatcgagtctcaacctcg	3474
Db	3240	TTTCATTAATGACTGGAAAGTGTGCCCTTCCTTCACCTGCTTATCGAGTCTCACCTCG	3299
QY	3475	taccgggtctgtccagagacaagcccatcaagttactccgggagacaggaagggccgaca	3534
Db	3300	CACGTGGCTGTGTCCAGACAGCAGTCGCTGAAGTACTCCCGTATACAGAGCAAGGGGCCGCA	3359
QY	3535	gcttccacagacttccacccaagaagaaacattlggaagaaagtgctatgata tagaggaag	3594
Db	3360	GCTTCTCTGACTGCCACCAAGGAACACTGMAAGAGTATGATGATATAGAGGAGAG	3419
QY	3555	aaatcgccaatatgaaacttaacaatacaaaacaggtagccgagatcaga cccccagactgga	3654
Db	3420	AAATGGCCAAATGAATCTTAACAAATACAAACAGGTAGCGGATACATATCCCAACATCGA	3479
QY	3655	gcaagattacacatlcgaagtalatccgcccagagatlgga tccacatagaaagggcaga tactgt	3714
Db	3480	ACAAATTTACCTTGTGAAGTATTCGCTCAGAGATGGGATCCACATAGAGGGGCAGATCTGT	3539
QY	3715	ttccaataatccctcggacaagtatg taaagtga tgtatctcggcgtttccaaggaatag tag	3774
Db	3540	TTCTACTAATCTCTGGAGACAGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGT	3599
QY	3775	tgcttctcglttcaagcagcacaaagctacatgtctccatccatcagagcgcgcgagaagaa	3834
Db	3600	TGCTTCTGCTTTACAGACAGCAAAAGTACATGTGCTGCTCAATCAGAAAGGCCGAGAGAAA	3659
QY	3835	caggagaataatcga tgcctttacatccaaatlgcaaaaacagacaaatlggcgtgtcgggaa	3894
Db	3660	CAGGAAATACATGTGTCTTACGTCCAAATATGCAAAAGCAGACAGCAATGGGTGTGTGGGGAA	3719
QY	3895	gaacttgaaccaaagacacacagatccag tlggaacatgtgtc tcaactggaagaa tgaacg	3954
Db	3720	GAGCATTTGGCCAAAGACACGATCAGTATGTGAGACATGTGCTCACTGGAGAAATATGTGG	3779
QY	3955	cagccagtcgcgaactgtcag tggcgccctggg taccag tlgcaagaagacggcgatc tag	4014
Db	3780	CAGCCAGTCGCGACACTGCGAGTGGGGCCCTGGGCGACACAGTGGCAAGAAAGGCGCATTTAG	3839
QY	4015	catlgggggccaaatgtgtatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	4074
Db	3840	CATTGGGGCCAAATGTGTACTATTTTGTGTCTCTCCGAAAAATGTGCAATGCTCTCCA	3899
QY	4075	actcagccaacaaagagagatgaaagctaaagctacgagagactgttccagaagaacacgga	4134
Db	3900	ACTCAGCCAAACCGAAGGAGAGGTAAAAAGCTAACGACGACTGTCCAGAGAAACACGGA	3959
QY	4135	gaccgggtacgagtggagatgaagaaactgtgatacccgccagggccagccgggaatccac	4194

Db	3960	GACCGGGCTAGCAGCTGACATGACATGAGCAACTGATGACAAAGCAGGCCGCGGGAATTCAC	4019
QY	4195	agaatgacacatgaacacagctatagctcggaaaggaanaalcgatcttccctggggtcgcct	4254
Db	4020	GGATGGACACATGAAACAGTTTACAGTTCAGAAAGGAATTCGATCTTTCCTGGGGTCCGCT	4079
QY	4255	ggcctctgacagccagttcaatgatattccttgatctggcctggccctgactagatggg	4314
Db	4080	GGCCTCTGACACCCCACTTTCAGCGATTTCCCTGTATGGTCTGTGGGCCCTGCTCAGCTAGTGGG	4139
QY	4315	agcgagaccctggcgtacccctgcacattgggttgcattcagtgaggaaatgtgtgtaaaaa	4374
Db	4140	AGCGCCAGACCCCTGGCCACTCTCTGCAATGGGTGACATTCAAGTGGGATGTGGATTAGAA	4199
QY	4375	ggagacagctgtaggttagaatcattccggcgcgccggccctgtgtglaaaccaggttccaa	4434
Db	4200	GGGACACGCTGGAGGTAGAGATATATCCGGGGCTGAGGGCTTGTAGTAAACAGGTTTCCA	4259
QY	4435	gacactgcccagacacglatatgtcaaggltatctctgttagaacaacggagctgtacatgcaa	4494
Db	4260	GACACTGCCAGACACCTATGATGTCAAGGTGTACCTGTTAGAACAAAGGAGTCTGCATAGGCAA	4319
QY	4495	aaagaaaacccaagggtgtagaanaagaccctggggcccttaccagagatctgttctc	4554
Db	4320	AAAGAAACCAAGGTGGCCAGAAAGACGCTGGAGCCCTGTATACGACGCTGTATCTTT	4379
QY	4555	cgaaggaagaccccccaaggaggaggtgttacaacatgtctcgtgggaatatatgtcgat	4614
Db	4380	CGAAGAGACACCCCCAGGGAGAGGTGTACAGATCATTTGTGTGGGAGATTAATGGCCCAT	4439
QY	4615	ggatcaacaatcccttataatggaggtgcccagatactcttagatgaactggaactatccaa	4674
Db	4440	GGATCAACAATCTCTTATGTGGAGTGGCCCGCAGATACCTGTAGATGAACGTGAACATATCAA	4499
QY	4675	catgtgtatgtgatgtgttcaaaactcttccctccctccctccttagtagatccaactcggc	4734
Db	4500	CATGTGAATTTGATGGTGTCAAACTCTTCCCTCTTCCCTCTGTAGTAACTCAACTTGTGC	4559
QY	4735	acctctgacaagaagagcttcccaatcgtctctcgtgaagtttacaacggactcttactc	4794
Db	4560	ACCTCTGACAAAGAAGGCTTCCCAATCGTCTCTGGAAAGTTCATCTGGACCTTCTTACTC	4619
QY	4795	tcgttcatatgacaactataaactgtgtgtcaacaacacagcagatatacaaaaaccagaaga	4854
Db	4620	TCGTTTCATAGCAACTATATAAAGCTGTGTGCACAAACACACGTTACAAAAAAACGAGAAGAA	4679
QY	4855	aacgcacagatgtgaagacccctcgttaacatcgtatgctgtgtgtgtgtcttcaagagccca	4914
Db	4680	ACCACCCAGAGTGG-GGACCCCTGGTCCACATGCTATGCTGTATGTTGTCTCCGGAGCCCA	4738
QY	4915	cgcttagaggataccaacagctcctgtgtctgtgtctcagaagaa-----tctacaactgtggc	4970
Db	4739	CCTTTTAGGGATTAAGAGGCCATCTCTGTGCTTTCAGAGAGAACCATCCGTGCACAGTGTGCC	4798
QY	4971	ctagacaagg 4980	
Db	4799	CTAGCAAGG 4808	
RESULT 3	AF199330		
LOCUS	AF199330	5592 bp	5MRNA
DEFINITION	Rattus norvegicus RIM2-5A (Rim2-5a)	mRNA	ROD
ACCESSION	AF199330		04-JUL-2000
KEYWORDS	AF199330.1	GI:8925873	
ORGANISM	Norway rat.		
SOURCE	Rattus norvegicus		
REFERENCE	Wang, Y., Sugita, S., and Sudhof, P. C.		
AUTHORS	1 (bases 1 to 5592)		


```
|||||
Db 1221 ACCATCAAGCAAGATCTGTATCTAGCTAGCTGCAATGAAACAGCATGCTA 1280
Oy 1381 ttcataatgaagaactcgaagagctcgaaggacaagtcttataccacaaggacctaaa 1440
Db 1281 TTCAATATGAAACACTCGAAGAGCTGAGGACAAAGTTCTTATCCACAAAGGACCAAA 1340
Oy 1441 tcatagtccctccaccccccgagagccctatacgccttgatagacagacatgagcg 1500
Db 1341 TCATATGCTCTCTTACCCCTTCGACAGAGCCCTATACCGCTCGATATGACCAAGCTGAAGCG 1400
Oy 1501 cgcctgacctccctacggaagaacagcacactatagatcccaagctctgctgtgaggaagaagaa 1560
Db 1401 TCCGACCTCCCTACGGAACACACACACACTAGATCCAGCTGCTGCTGTATAGAAACGAA 1460
Oy 1561 gcgagaaaaaatggaacaacatgtlaagaaatgattctcttgatgcagacagtcgagtc 1620
Db 1461 GCGAATAAAATGGAACAACTGTGAGGAATGATCTTTCATGTTCAACACAGCTGTGAGTC 1520
Oy 1621 agtgaagccgcccccaagaagccctataatccaaagaagaaggtataaagtgcagagtc 1680
Db 1521 AGTGAAGCCGCCCCCAACAGCCCTCTAATATCCAAAGAAAGAGGTAAATGCCCAAGCT 1580
Oy 1681 ttcactgaagcaagctcgaagagagagcttgagcatccacacactgaatatacaagctgtga 1740
Db 1581 TTCACCTGACAGCTCAGAGAGAGAGAGTGGCATCCAGCCTGAGATATCAAGCTGTGATGA 1640
Oy 1741 tctgaagctcggaagaagcgaagtgatgtagtgaagaaggagacatgcaaaagggaagaaa 1800
Db 1641 CCGTGAGATTGAAAGCGAAGCGTAGCTGAGAAAGGAGAA----- 1681
Oy 1801 aaactagtgagcaggaagtttgcgagatcttaacacagagctctgagagacaaagaagaag 1860
Db 1682 ----- 1681
Oy 1861 gatgtactatggttgccactcttctggaagaagaattggaatggtctgagctcagatataa 1920
Db 1682 -----TGAGTACAGCTGGTTGAGCATGCGCTCTTG 1712
Oy 1921 ggaactcgaggtatagataactgtgtagcacacaaccttaacgaagagacatagatga 1980
Db 1713 G-----CATAGCAGTGCAGCANTCCCAATGTC 1739
Oy 1981 taagcaacctgtgacctggaagcaccatccaaagaatgagatcgaccttaattgctatttt 2040
Db 1740 TTTCGACCTGTGACCTGGCAGCCATCCAAAGATGGAGATGCGCTTAATGGCTCTAATTTT 1799
Oy 2041 attaaataagcgtltaaaagaatlgagagtgctacctcgaagactcgaagcaatgctgggctt 2100
Db 1800 ATTAAATAGCGTTTAAAGATGGAGTGTACTCGATCGAGATTCAAGGACCAATGCTGGGCTT 1859
Oy 2101 aaaggtctgtaggaagaagaatgactgaatcagaagtcgaacttgatcatlataccaagt 2160
Db 1860 AAAGGTTGTAGGGGGAAGAAATGACTGATCATGAGTGCAGCTTGTGCAATTATTTACCAAGT 1919
Oy 2161 aaaaaaagaagaatttagctgatactctgtagacatcttaagccaggtatataaagcttggga 2220
Db 1920 GAAAAAAGGAAGTTTAGCTGATACGTAGAGACATCTTAGACCAAGATGAGAGCTTGGGA 1979
Oy 2221 atggaatggaaggtctatctgaagaagccacaattggaagaagtttacaacatlatctaga 2280
Db 1980 ATNGAATGGAAGGCTATTGCAAGAGAGCAATTTGAGGAAGTTTACAACATTTATTTAGA 2039
Oy 2281 atccaaactgaacaacaagatgtagctgtgtcttccaaggccaattgagagatatctctag 2340
Db 2040 ATCCAGAGCTGGAACCAAGATTGAGCTTGTGTTTCAAGGCCGATGTGAGATATGCTAG 2099
Oy 2341 aatactctgtagcagcagacagacaacttgaaatccaggtctagctcatltaaatccaana 2400
Db 2100 AATTACTGTATGACACTCAATGACACAACTGGAACTGAGTTCTAGCTCATTTGATATCAAAA 2159
Oy 2401 aatgagacgtctctctatactcglttaacctaaccaatgagtcgagacgtctgaggaagt 2460
|||||
Db 2160 AATGATGCTCCGCTATATCGGTTACTCTCCCATGAGTCTGGCATGCTGAGGAGAT 2219
Oy 2461 cccgagttctatctcgtgagcagcttcaataaactatggttgaagaagttgtaacca 2520
Db 2220 CCCACAGTCTTATCTGGACAGCTTTCATTAATAAATATGATGGTTGACAAAGGTTGCTCACCA 2279
Oy 2521 gttgtagttacaatttlttggaagaaagatctcccttccaaggaaagatgagagccaag 2580
Db 2280 GTTATATGATTACAATTTTGGACACAAAGAGATCTCCCTCCAGGGAAAGATGGAGGCCAAG 2339
Oy 2581 gaatccctatgtlaagaattacttccctccagaatagaatgataaataagaagaagaac 2640
Db 2340 GAATCCTTATGTTTAAATTTTACTTCTTCACAGACAAGATGATTAATAAACAAGAAAGAAC 2399
Oy 2641 aaaaacagtcgaagaacttltggaacccaatggaaccagacttcaatttctctctgt 2700
Db 2400 AAAAACAGTCAAGAAAACTTTGGAACCCAAATGGAACCAACACTTTCATTTATCTCTGT 2459
Oy 2701 ccacggaagagaatcccgtyaacgaaatgctggaatltacccctltggatcaagaatagat 2760
Db 2460 TCACCGAAGAGATTCGGGGAACGAATGCTCGAATTCACCCCTTTGGATACAGCTGAGAT 2519
Oy 2761 tcgagaagaagaagcgaattcctlaagagaatlttaattgaattggaacagcttgc 2820
Db 2520 TCGAGAGGAAGAAAGATTCCTAGAGAGATTTTAATGAAATTGGAACAGCTCTGT 2579
Oy 2821 agatgatgagccgcaactgtgtataagctgcagaccatgattgtctctcatctgcaccc 2880
Db 2580 AATATGATGAACACACTGATACAACTTCAGACCCATGATGTCCTCATTTGCCACTGCC 2639
Oy 2881 tcgcccctcccatctcgcggagagcagctccatctgagagaagcccaagcgaagct 2940
Db 2640 TCACCTTCTCCATATATATCCACAGAGACAGCTGCAATGAGAGAGCCCAACACGACAGTT 2699
Oy 2941 gcaaaagtcgaagaagaataagtgacagtgtaagtgtctgactacagactcgaagatggc 3000
Db 2700 GCAAAAGGTGGAAGAGATTAAGTACAGAGAGATGATCGAATGATGACGAGAGCGGCT 2759
Oy 3001 gggagtgaagtcaagattatcgcacacaatgycggagcttctcaaaagctccagctgtgt 3060
Db 2760 GGGAGTCGTGAGATTACCGACACGATGGCCGTGATCTTCAAAAGCTCCACATTTATAGT 2819
Oy 3061 gccagaacaagtcattgcatcaaatcatgtctcacaatcaaggtctcctcatcagataga 3120
Db 2820 GCCAGAACAGTCAATCTCATCAAAACCATTCCTACCAATCAGCTCTCTATGAGATAGA 2879
Oy 3121 tgtataagaaggaacaaggtlcatgtgcagctagtgccctcctcctcaaaagaatgtgga 3180
Db 2880 TGTATATAGGAAGACTAAGTTCATGTGTCGCTAGTGTCTCTCCCTCAAAAGAAATGAGGA 2939
Oy 3181 acaagggaacccgaaggaacacgtgtactcgtgcattacaacaaattagccgaatgatatag 3240
Db 2940 ACAGGGGCTTCAGAGGAGACAGTGTCTACTGGCATTTCAAAATACAATTAAGCCGATGAGAG 2999
Oy 3241 acaacgtgtcatgtagtgaactactcttcaatagatagaag----- 3280
Db 3000 ACACCGGTGATGAGACGACCATTTACTCTTCAGAGAGACAGTCATTTTCTCACTATACC 3039
Oy 3281 -----agggaatgtgaagc 3294
Db 3060 TCGCTCTGACACAGGACAGACAGTAGATCACCAAGAGGATGGAAGGATTTGTGAAGC 3119
Oy 3295 agcagatagaacagccatatacagaatccagaatcaaacagaacacagcctctcctagaagcg 3354
Db 3120 AGCAGATTAGACAGCCTATATCACAGATCCAGATATCAAGAAACAACGGGCTTCTCTAGAGCG 3179
Oy 3355 gacacacacccgctccagatctcctctgaacgtccctgatacaaacctcaatgaggtcagatcc 3414
Db 3180 GACCAACACCCGCTCCAGATCTCTGAAGCTGCTGATACAAACTCTATGAGGTGATGCC 3299
Oy 3415 ttcattaatgactggaagaatctgcccctccttcaactgtccttcaagatctcaacctcg 3474
Db 3240 TTCATTATATGACTGGAAGATCTGCCCTCTCTACCGTCTTATCGAGGTCTCACCTCG 3299
```


SGMKHQAISDMPSDKRRSPSVSRDQNRRYDQSEEREYSQYVPSDSTMPSVDADR
 RSOREPOFYEEDPHLNTYRDSNRGRHRSKEYIVDEDEVSREBEYEQREBEYOARSR
 SDPNLARYPEPOPEEDMORSIMPLAEYSRAHERHRSVLSLANELESRIISLMDPS
 RORVSERRAMENORASYSMERTRAQCOSSYPORNTNINSVPBRPPRPHKSKGKMKR
 ADSLRKHQHLDPSSAVKRTKREKMETMLRNDLSLSDQSSVPRPPRPHKSKGKMKR
 QVSLSSSEELASTPEYVSCDVELESEBVGKDMEXWLSLHSAHSEASMSLAP
 VTWQPSKDGRLIGRLINKRLKLDGSPVSGAMGLKVVGKDMEXWLSLHSAHSEASMSLAP
 KGSADLVYGHLPBGDEVLEMMNRLLQAGFEVYNILLESKEPEYELVSRIGDMP
 RIPSTHQLSSSSSEFOSKMDRPSIVTSPMSGMLDVQOLESFSGOLSSLSLSTRT
 TPFPVROIKLMPDVGHOLVYTLIGARDLPREBGRPNPYKITYFLDRSRKMKR
 TKYKTLLEPKMNQFFITSPVHREFREMLITLMDQARVEESESFEIGELILELET
 ALLDDEPMWYLIQTHDVSSLPHPSPYMPRQLHESPTRLRQNRKRLSDSEVSDYD
 CEDGVSVYDRHGRDLOSLSTLYPEVYMSNHCSPPSGSPHVDYIGTRSRPSVP
 PPORNVEGLRGTNRATGHYNTISRMDRHYMDHNSSEDSHFLTLPRSRHOTSSEH
 HRDRCDEADROPHYHRSRTEORPLLETTTTRSSSERADPNLMSMPSLMTGRAP
 PSLALSRHPRTGSYQTPSSPTPYGRKROLPOLPKTKLEMTLEDMDSTRKRSNG
 AMDLERNRQMLNKTKOVASDPRLEODYHSKRGMDPHRADVTSRSDSDVS
 VSAVSRSSASRFSSTSYMASVOSEBRGRNRKISVTSKMSRQMSGSKSMARSTIS
 GDMCSLEKNDQSODTAVGALGTSGKRRSSIGAKMVAIVLSRQMSRQMSOLSTEG
 GKRLSTVORSTETGLAEMRMMTRQASRESTDGSNYSBGNLIFPGVRLASDQ
 FSDFLDGPPOIVOROTLAPPMAGDIOVGMMDKKGOLFEVETIRAGLIVYKPSKITP
 APYKVTLLDNGVCIARKKTKYARKTLERLYOOLISFEESPOGRVQIITYWQYGRMD
 HKSTMGVAQIILDELISNMVIGMFKLPPSSLVDFTLAPLTRRASQSSLESSTGPSY
 SRS"

BASE COUNT 1695 a 1301 c 1409 g 1235 t
 ORIGIN

Query Match 75.8%; Score 3776; DB 94; Length 5640;
 Best Local Similarity 86.4%; Pred. No. 0;
 Matches 4443; Conservative 0; Mismatches 275; Indels 422; Gaps 8;

OY 1 gcttcctcagggtggttgcgtccccaacatgtgcgtccgctcgggcccggggccg 60
 |||||||
 DB 21 gcttcctcagggtggttgcgtccccaacatgtgcgtccgctcgggcccggggccg 80
 |||||||
 OY 61 cccggctcccaacccggcggtctcgaacctccctccgagccgagatgcccag 120
 |||||||
 DB 81 cccggctcccaacccggcggtctcgaacctccctccgagccgagatgcccag 140
 |||||||
 OY 121 ccaactcaggaagaagaagaataatcactgctgctgcatatgcatcgtcagaaga 180
 |||||||
 DB 141 ccaactcaggaagaagaagaataatcactgctgctgcatatgcatcgtcagaaga 200
 |||||||
 OY 181 agagagaagaagaagcgtccgtgctcagaatcagaagaacacacaacccagaca 240
 |||||||
 DB 201 agagagaagaagaagcgtccgtgctcagaatcagaagaacacacaacccagaca 227
 |||||||
 OY 241 gtygttcctccttagtggatcactgaactggttaatacgttctcgagcccgacaaa 300
 |||||||
 DB 228 gtygttcctccttagtggatcactgaactggttaatacgttctcgagcccgacaaa 227
 |||||||
 OY 301 acaaccaatgagaagaagccccaagaacatgcaaccaaatltgtaaatgtataaaga 360
 |||||||
 DB 228 acaaccaatgagaagaagccccaagaacatgcaaccaaatltgtaaatgtataaaga 260
 |||||||
 OY 361 gcaagtcaagaagaatggagagaagaatcgcagacagcagaagaagtgatgcccc 420
 |||||||
 DB 261 gcaagtcaagaagaatggagagaagaatcgcagacagcagaagaagtgatgcccc 320
 |||||||
 OY 421 gacctgtgcatctgcacaagaacaaatttgcagatggaatgagcggccatattgttcta 480
 |||||||
 DB 321 gacctgtgcatctgcacaagaacaaatttgcagatggaatgagcggccatattgttcta 380
 |||||||
 OY 481 ttgcgaacaaagttctgtgctcgatgtggaaggtcgagtgtcttaacgtcaaaacaggt 540
 |||||||
 DB 381 ttgcgaacaaagttctgtgctcgatgtggaaggtcgagtgtcttaacgtcaaaacaggt 440
 |||||||
 OY 541 tatgtgggtgtgtaatttggcgaacaaacagaagaatctctcaataatcaagagatg 600
 |||||||
 DB 441 tatgtgggtgtgtaatttggcgaacaaacagaagaatctctcaataatcaagagatg 500
 |||||||
 OY 601 gtttataatagtggtgtctaacaactgcagcaactgttcaaaaggttcctcgtagagct 660
 |||||||

|||||
 DB 501 gttttacaatactgaggttaataacaccgacgacctgatcaaaaaggctcttcgagagct 560
 |||||||
 OY 661 tcgaatatgagaagcccccccaagagaagaagaacaaactacacagcagcccaagttcca 720
 |||||||
 DB 561 tcgcaatatgagaagcccccccaagagaagaagaacaaactacacagcagcccaagttcca 620
 |||||||
 OY 721 aggaagccccaggtgacttaccagcttcaggttgaagaagcgcagcctatgaggtcac 780
 |||||||
 DB 621 aggaagccccaggtgacttaccagcttcaggttgaagaagcgcagcctatgaggtcac 680
 |||||||
 OY 781 aagacagatatactataaaaatggaatcagaagatgaaacccaatltgcagtgaatgac 840
 |||||||
 DB 681 aagacagatatactataaaaatggaatcagaagatgaaacccaatltgcagtgaatgac 740
 |||||||
 OY 841 ttcaagacagaanaaagaatccatcagatgtcccaaggaatcgaatcgaagaatagaaga 900
 |||||||
 DB 741 ttcaagacagaanaaagaatccatcagatgtcccaaggaatcgaatcgaagaatagaaga 800
 |||||||
 OY 901 tgaagaagaagaagactactcaacagatgttccttcaagatggtacaaatgccaagatctcc 960
 |||||||
 DB 801 tgaagaagaagaagactactcaacagatgttccttcaagatggtacaaatgccaagatctcc 860
 |||||||
 OY 961 ttccgattatgctgtagagacgaatcctcagcgtgagccccaatlttataagaagaactggtca 1020
 |||||||
 DB 861 atcagattatgctgtagagacgaatcctcagcgtgagccccaatlttataagaagaactggtca 920
 |||||||
 OY 1021 tttaaatcaagaggatcttaacagaagaagccaatagaacatltccaagaagatattgttga 1080
 |||||||
 DB 921 tttaaatcaagaggatcttaacagaagaagccaatagaacatltccaagaagatattgttga 980
 |||||||
 OY 1081 tgatgaagatgtyggagaagaagatgaatatagaagaacaaagaagaagaagaataatcca 1140
 |||||||
 DB 981 tgatgaagatgtyggagaagaagatgaatatagaagaacaaagaagaagaagaataatcca 1040
 |||||||
 OY 1141 ggcagcgtcaacaagaatgtaaccgaatctggccggtatcccgtaaaagccaacactcaga 1200
 |||||||
 DB 1041 ggcagcgtcaacaagaatgtaaccgaatctggccggtatcccgtaaaagccaacactcaga 1100
 |||||||
 OY 1201 agaacaatatgcataccacgcgtggaaggtgtccaggaagaacatgtagagaagaagcagtgta 1260
 |||||||
 DB 1101 agaacaatatgcataccacgcgtggaaggtgtgtccaggaagaacatgtagagaagaagcagtgta 1160
 |||||||
 OY 1261 gtttctcttggcaaacgctggaactagaagaatltccaggaatltctctgctaagaagatgtag 1320
 |||||||
 DB 1161 gtttctcttggcaaacgctggaactagaagaatltccaggaatltctctgctaagaagatgtag 1220
 |||||||
 OY 1321 accatcaaggaagaagaatgtaactgtatctgaacgttagagctgcaatggaagaacacagatcgtta 1380
 |||||||
 DB 1221 accatcaaggaagaagaatgtaactgtatctgaacgttagagctgcaatggaagaacacagatcgtta 1280
 |||||||
 OY 1381 ttcaatggaagaagaactcgagaaggtcgaaggaacaaagtltcttataccaagaagaactcaaa 1440
 |||||||
 DB 1281 ttcaatggaagaagaactcgagaaggtcgaaggaacaaagtltcttataccaagaagaactcaaa 1340
 |||||||
 OY 1441 tcatagctctcccaacccctcgagcggagccctataccgcttgaatagaaccagaactgaagcg 1500
 |||||||
 DB 1341 tcatagctctcccaacccctcgagcggagccctataccgcttgaatagaaccagaactgaagcg 1400
 |||||||
 OY 1501 cgtctactccctcaggaagaacacgaacacttagatltcccaagctcgtctgttggagaacaaagaa 1560
 |||||||
 DB 1401 cgtctactccctcaggaagaacacgaacacttagatltcccaagctcgtctgttggagaacaaagaa 1460
 |||||||
 OY 1561 gcgagaanaaaatggaacaaatgtaggaatgtagtcttgaaggtcgaagcagtcgagtc 1620
 |||||||
 DB 1461 gcgagaanaaaatggaacaaatgtaggaatgtagtcttgaaggtcgaagcagtcgagtc 1520
 |||||||
 OY 1621 agttagagccgcccccaagaagcctcaataatccaagaagaaggttaaaatgtgcccaggt 1680
 |||||||
 DB 1521 agttagagccgcccccaagaagcctcaataatccaagaagaaggttaaaatgtgcccaggt 1580
 |||||||
 OY 1681 ttcaactgagcagctcgaaggaagagctggtgcatcacaactgtaatatacaagcgtgttga 1740
 |||||||

Db 1581 TTCACCTGACGAGCTCAGAGAGAGATTGGACATCCAGCCTGAGTATACAGCTGTGATGA 1640
QY 1741 tctggagctctgaaagcagagtgctgagtgaaagaagggacgctcaaaagggaaaaaaga 1800
Db 1641 CGTGGAAATTGAAAGCCAGAGCGTAGGTGAGAAAGAGAGACAT----- 1682
QY 1801 aactagtgagcagggagtttctgcgattctcaaccagjctcgagagcaaaagaanaag 1860
Db 1683 ----- 1682
QY 1861 gatctactatgtyggccaactcttggaaggagalttgaaatgylctcgacctcagatkaa 1920
Db 1683 -----GGAGTACAGCTGGTGGATGACATGCGCTCTTG 1712
QY 1921 ggaactcggggtagatacctgtagtagcacaaccttaagcagagagcatagcatalgta 1980
Db 1713 G-----CATAGCAGTGGAGCATCCCAATGTC 1739
QY 1981 taagcacacctgtgacctgagccatccaaagatggagatcgccaatctgtcgctattt 2040
Db 1740 TTTGCAACCCGTGACCTGGCAGCCATCCAAAGATGGAGATGCCCTTAATTGGCTGATTTT 1799
QY 2041 attaataagcgtttaaagaatgggagtgtaactcctgagactcaggaagcaatgctgagcct 2100
Db 1800 ATTAAATTAAGCGTTTAAAGATGGAGTGTACTCGAGATTACAGAGCAATGCTGGCCTT 1859
QY 2101 aaaggttgtgagggaaaagatgactgataagtcgactttgtgcatltattaccaagt 2160
Db 1860 AAMGTTGTGAGAGAAAGATGACTGAATGAGTGAGTGTGCTGATTTATTAACAAAGT 1919
QY 2161 aaaaaaagaagtttagctgatactgtgtagaacaatcagaccagtgatgaagtcctgga 2220
Db 1920 GAAAAAGGAAGTTTACGTGATCTAGTGTAGGACATCTTAGCAGAGTGATGAAGTCTTGA 1979
QY 2221 atggaaatggagagctatctgaaagagcccaatttgagggaagtttaacaataltctaga 2280
Db 1980 ATGGAATGGAAGGCTATTGTGAAGAGGCCAATTGGAGAACTTTAACACATTATCTCTGA 2039
QY 2281 atccaaaccctgaaaccaagaattgagctgtgtgttcaaggccaattggagatatccctag 2340
Db 2040 ATCCAAAGCCCGAAGCCACAAATGTGAGCTTGTGTTCAAGGCCGATTGGAGATATGCCCTAG 2099
QY 2341 aataactgtagaagcaatgacaaactggaatccagttctagctcaatttgatctcaaaa 2400
Db 2100 AATACCTGATAGACATCATGACAACTGGAATCCAGTTCTAGTCATTTGATTCATAAA 2159
QY 2401 aatggacgcgtctctatataccgcttactcaaccaatgagtcctgagcatcgaggagatgt 2460
Db 2160 AATGGATCGCTCCGTATATCGGTACCTCTCCATGATGATGCTGGCATGCTGAGGATGT 2219
QY 2461 ccgcgagctcttactctggaagacttca----- 2488
Db 2220 CCCACAGTCTTATCTGAGACAGCTTTCAAGCCAAAGCCTTAGTGAAGAACAGCGCTTT 2279
QY 2489 -----ataaactatggttgaacaaggttgctacacagttgatatctac 2532
Db 2280 TGTTCCTAGAGGTTTCAGATTAATACTATGTGTTGACAAAGTTGGTCAACGATTAATAGTTAC 2339
QY 2533 aatttgggaagcaaaagatctccctccagggagaatggyagggccaaggaatccctatgt 2592
Db 2340 AATTTTGGAGCAAAAGATCTCCCTTCCAGGAAAGATGGGGCCCAAGGAATCTCTTATGT 2399
QY 2593 taagattactctctccagatagaagtgaataaataagaagaagaacaaaacagttcaa 2652
Db 2400 TAAATTTTACTTCTTCCAGACAGAAATGATTAATAACAAAGAGAAACAAAACAGTAA 2459
QY 2653 gaaaacttggaaacccaatggaacagagattcaattatctcctctccacgaaagga 2712
Db 2460 GAAACTTTTGAACCCAAATGGAACAGACTTTCATTTATTTCTCTGTTCCCGAAGAGA 2519
QY 2713 attcogtgaagaaatgctggaatattacacttgggatcaagtagagttcogaagaaga 2772
Db 2520 ATTCCGGGAACGAATGCTGAAATCAACCCCTTTGGGATCAAGCTCGAGTTCGAGAGAAGA 2579

QY 2773 gacgcgaattcttagggagagatttaattgaattggaacagacttctgtagatgagcc 2832
Db 2580 AAGTGAAATTCCTAGGAGAGATTTTAAATTGAATTTGAAACACACTCTGTATGATGAACC 2639
QY 2833 gcactgtataaagcttgagagcccaatgactctctcaattggccactccctccctccc 2892
Db 2640 ACACCTGTATCAAACTTCAGACCCATGATGTCTCTCATTTGGCACTCCCTTCTCTCC 2699
QY 2893 atactgccccggagggacatccatggaagagcccaagcgcaagctgcgaaggtcgaa 2952
Db 2700 ATATATGCCACGAGAGACGCTGCATGAGAGAGSCCAACAGCGATTTGCAAGAGTGA 2759
QY 2953 gagaataagtgagagagatgactgactcagactcgagagatgagctggagatgactgc 3012
Db 2760 GAGAAATTAAGTGAACGAGCAATATCCGACTGTGACTGGAGAGCGCTGGAGTGTGTC 2819
QY 3013 agattatcgacacaatgagccgcatctcgaagctccagctgtgcggtgccaagacaagt 3072
Db 2820 AGATTACCGACAGATGGCCGTGATCTTCAAAAGCTCCACATTATCAGTCCAGAACAA 2879
QY 3073 catgcatcaaatcactgctcaccatcaaggtctcctcattcgaatgagatgatalaggaag 3132
Db 2880 CATGTATCAAAACCATTTGCTACCATCAGGCTCTCCATCGAATGATGATTATAGGAAG 2939
QY 3133 gaaagtgatggtgctgactgagccctcctcctcaaggaatgtygaaicagggacag 3192
Db 2940 GACTAGTGTATGTGTGCGCTAGTGTCTCTCTCTCAAGAAATGTGAAACAGGGGCTTGC 2999
QY 3193 agggacagctgctactgagccattcaacacaatgaacagatgtagaagacgctgcat 3252
Db 3000 AGGGACACGTGCTATGCGCATTTACAAATTAATGAGCGGATGACAGACACCTGTGAT 3059
QY 3253 ggaatccactactcttcagatagagac----- 3280
Db 3060 GGAGACCAATTACTCTTCAAGAGAGACAGTCAATTTCTCACTACCTGCTCTCGACA 3119
QY 3281 -----aggagttggaagcaggaatgaca 3306
Db 3120 CAGGCAGACCATGAGCATACCAACAGGATGGAAGGATTTGTAACGACGATGATGACA 3179
QY 3307 gccatactcaagatccagatcaacagaacaagcgccctcctctagagcggaaccaaccgg 3366
Db 3180 GCCATATCAAGATTCAGATTCACAAAGAACAGAACGCGCTCTCTAGAGGGACACACCCG 3239
QY 3367 ctccagatcccttgaaagctctgatatcaaacctcaatgaggtcgatgctcctcaatgac 3426
Db 3240 CTCAGATCTCTGAAACGTGCTGATCAAAACCTCATGAGGTGCGATGCTTCATTAATGAC 3299
QY 3427 tggaaatctgcccctccttcaactccttaatgaaagttccacactgtaocgggtctgt 3486
Db 3300 TGGAAATCTGCCCCCTCTCTCACCTCTTATGAGAGTCTCACCCCTGCACTGGCTGTGT 3359
QY 3487 ccagaaagcccaatcaagatctccggaacaggaagggccgaagcttccacagct 3546
Db 3360 CCAGACCACTCGTCAAGTACTCCGTAACAGSACGAAGSGCGCGGACGTTCTCTACGT 3419
QY 3547 tccaccaagggaaacttggagagaa----- 3572
Db 3420 CCACCCAAAGGGAAACACTGGAAGAAATGATCACAGAAAGATGAGCTCCACAAGAGAG 3479
QY 3573 -----gtgctatgatatagaggagaagaatcgccaatgaaacttaacaataca 3624
Db 3480 AAATCTAGGTGCTATGATATAGAGAGAAATTCGCCAAATGAATAACTTAACAATTCAA 3539
QY 3625 acaagtgacggatcagaccacagactggaagaaatattacattcgaagratcgctcag 3684
Db 3540 ACAGGTAGCGGATGATCGATCCAGACTGGAACAAAGATTACATTTCAAGATATGCTTAGG 3599
QY 3685 atggatccacatagaagggcagatactgtttccactlaaattcctcgagacagtatgtaa 3744
Db 3600 ATGGATCCACATAGAGGGGAGATGACTGTTCTTACTAATATCTCGGACAGTATGTAAG 3699

QY	3745	tgatgtatcttgcggttccaagactagtagtgctctctcgttccaagcacaagctacat	3804
Db	3660	tgatgtatcttgcggttccaagcactgtagtgcgttcttcgttcagcgcacaaactctacat	3719
QY	3805	gtccgttccaatagaacccgcgcagagagaacacagaaatacagtgcttttcatccaatat	3864
Db	3720	gttcctgccatatagaacggcccgagagaaacagaaatatagtgcttcttcatccaatat	3779
QY	3865	gcaaaacagacagatbgygcgtlctggggaagaacttgaaccaaaagccacagactcagtg	3924
Db	3780	gcAAAGcAGAcAGATGGGtGtGTCGGGGAAAGACnTGGCCAAAGcACAGATCAATGG	3839
QY	3925	agacatgtgtcactctgagagaagaatgaacggcaacagctccgcacatgtagtggccct	3984
Db	3840	AGACATGTGCTCTACTGGAGAAGATATGTCACCGTCGCGACACTCAGTGGCGCCT	3899
QY	3985	gggttccagatggcaaaaacgagatctagactctggggccaataatgtaactatgtgtg	4044
Db	3900	GGGcACcAcATGGcAAAGcGGGATCTAGcATTGGGGCCAAATGGTACTATTGTTGG	3959
QY	4045	tctctcaagaaaagctgcagtgccctcactcaactcaacccaacccaagagagagtaaaaa	4104
Db	3960	TCTCTcAGGGAAAAGTCGcAGTGCCTCCCACTCAGCGAAACCGAAGAGAGGTAAAA	4019
QY	4105	gctaagagacactgttcaagaagaacacgagacgggctagcagtgtagatbgaagaactg	4164
Db	4020	CTTACGcACAcACTGTCTCAAGAGAAGcACGGcAGCCGGCTAGcAGTGAGATGAAGAcTG	4079
QY	4165	gltgcacccgcagggccagcgcgggaatccaacagtlgtcagatatgaacgctatagctcga	4224
Db	4080	GAATGcAAAGcAGCCAGCCGGGAATTCACAGATGGcAGATGAACGTTTACACTCAGA	4139
QY	4225	aggaatctgcactctccctgggggtcccgctgycctctgcacagcagtlcagtgatctcct	4284
Db	4140	AGGAATCTGATCTTTCTCTGGGGTCCGGCTCGGCTCTGACAGCGAGTTCAAGCATTTCT	4199
QY	4285	ggaatggcctggccctgtcgaactagltgggacgcacagccctgtgctactcctgtcaatgg	4344
Db	4200	TGATGGTCTGGGCCCTGCTCAGCTAGTGGAACCGcAGCCCTGGCGACTCCTGCAATGGG	4259
QY	4345	tgaacttaaggtgggaatgatatgaataaaaggaacagctbgyaagttagaatacatccggc	4404
Db	4260	TCACATTCAGGTGGGATGATGATGAAGGGACAGcAGCTGGAGTAAAGATCAATCCGGC	4319
QY	4405	ggcggcctctgtgtaaaaacgaagtccaagaacatgcagacacgtatbtaagtgta	4464
Db	4320	TGGAGGCTCTGTAGTAAACACAGTTTCCAAGACACTCCAGCACATATGTCAAGGTATA	4379
QY	4465	tctgttagaacaacggaactgtcgaactgaaccaaaagaaaccaaagtbgcggaagaacct	4524
Db	4380	CCTGTAGACAAcAGGAGTGTGCTATGCCCCAAAAGAAACCAAGGTGGCGAAAGACGT	4439
QY	4525	ggagccctctgtaaccagcaactctgttcccttcggggagagcccccagggggaggtgttaca	4584
Db	4440	GGAGCCCCGTATCACACACTGTTATCTTTGGAAGAGAGCCCCCAGGGGAAGGTGTACA	4499
QY	4585	gatactgtctggggagatatgtgcgatbgaatcaacaaatcccttlatbgyagtgccca	4644
Db	4500	GATCATGTGCTGGGGAGATTATGGCCGCAGTGAATCAAAATTCCTTTATGGGAGTGGCCCA	4559
QY	4645	gatactctagaatgaactbgyaactatccaacatcgaatgtagtbtgaatbgttcaaatctccc	4704
Db	4560	GATACTCTTAGATGAACCTGGAACATATCCAAcATGTGTAATGGATGGTTCAAACTCTTCCC	4619
QY	4705	tccctctccctcagtagataccaactccggaactctgacaagaagagcttcccaatcgtc	4764
Db	4620	TCTCTTCTCTCCAGTAGATCAACACTTGGCACTCTGACAAAGAGACCTTCCCAATGCTC	4679
QY	4765	tctggaaagtctaacccggaccttcttactctgttcaatgaactataaaactbgttca	4824
Db	4680	TCTGGAAGTCTTACTAGGACTCTTACTCTGCTTCACTAGCAACTATTAANACTTGTGCA	4739
QY	4825	caacaacacagcatalcaaaaaacagaagaacacgcacagltggaagccctgttgaacct	4884

[illegible]

BASE COUNT 1555 a 1190 c 1286 g 1141 t
ORIGIN

Query Match 64.2% Score 3197.6; DB 94; Length 5172;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 4101; Conservative 0; Mismatches 239; Indels 710; Gaps 5;

1 gcttccttaagggtggttcgcgtccgcaccaacatgctcgcgtccgcgtcgggcccgggcg 60
|||||
21 GCTTCCTTAGGGGTGGTGGCTCCACCAACATGTCGGCTCAGTCGCCGCCCGGGGCGG 80
61 ccgggtcccaaccccgcgcgccctccaacctctccgcagcccgagatgcccgaactcg 120
|||||
81 CCCGGCTCCACCCCGGGGGCTCTCAGCCTCTCCGCGCCGAGATCCGGACCTCAG 140
121 ccactccaggaagagagagaataatcatctgtgctgcatgagatcgctcagaagaaga 180
|||||
141 CCACCTCCAGGAGAGAGAGAAATCATCCAGGCTGTCAATGATGATCGTCAGAAAGA 200
181 agagagaagagagcagtcgctgctcaagaatcaagaagaacaacaagcacaaccgacaca 240
|||||
201 AGAGAGAGAGAGCAGTCCGTCTCAA----- 227
241 gtggttcccttagtgggatacctgaactggttaataaagttctgcagcccgcaaaa 300
228 ----- 227

301 acaacccaatgagaagagccccaagacaagctgcaaccaaatltgaaatgtataaga 360
|||||
228 -----AAAGCTGCATCAACAMTTTGAATGTATAAGA 260

361 gcaagtcagaagaatgaggagagaatcgcaagcagcagaagcagaaggtatgcccc 420
|||||
261 GCAAGTCAAGAAATGGGAGAGAAATCACAAACAGCAGACAGAGGCGCAGGCC 320
421 gacctgtgcatctgcacaagaacaatttgcagatgagatgctgcgcataattgctcta 480
|||||
321 GACCTGCGGCATCTGCCACAGACAAATTTGCAGATGATGGCGCATACGTGTATA 380
481 ttgcaaaccaagttctgtctcgatgctgaggtcgaggtcagtgcttcaagctcaaacaggt 540
|||||
381 TTGCCAAACCAAGTCTGTGCTGTGTGGAGGTGAGTGTCAATTACGCTCAACAAAGGT 440
541 tatgtggtgttaatttctgcccgaacaacaagaanaattcctaactaaatcaggaagt 600
|||||
441 TATGTGGGTGTAAATTTGTGCCGAAACCAAGAAATCTCTCACTAAGTCGGCGGTG 500
601 gtttataatagtggtctaacacactgcagcaacctgatacaaaagttccctcgaggct 660
|||||
501 GTTTTTCATATAGTGGGTCTAATACACCGCAGAGCCTGATCAAAAGGCTCTTTCAGGGCT 560
661 tcgaaatgagaagccctcctcagaagaagaacaacatacagagcagccccagttcca 720
|||||
561 TCGCAGATGAGGAAGCCCTCAGAGAGAAAGCAAACTCATGAGCAGAGCAGATTCCA 620
721 agagagcccaagtgactatcaactacgtgaattgagaagaagccgaactcaatgggttac 780
|||||
621 GGGACCCCGGGAGCTCATCACTACTGAGTGGAGAGAGTGAAGTCTATGGGCTAC 680
781 aagacagagatactataaataatgatcagaagtgaaagcacacagatltccagtgacatgc 840
|||||
681 AAGACAGGATTTCTATTAAAGATGGCTCAGGAATGGAACACAGATCCGACATGCC 740
841 ttcagacagaaacgaagttccatcagtgctcagggatcaaaatcgaagatcagcgcaag 900
|||||
741 TTCAGACAAAGAAAGAGTCCGTAGTCCAGGATCAAAATGGAATVCGACCAAG 800
901 tgaagaagaagagagactcacagiatgttccctcagatgtagtaaatgcaaatctcc 960
|||||
801 TGAAGAAAGAGAGATATTCACAGTATGTTCTTCAATGACAAATGCTAGATCTCC 860
961 ttcgagttatgtcatagacatctcagcgtgagccctcaatttatgaagaacctgttca 1020

|||||
861 ATCAGATTATGCTATGAGCATCTCAACGTAACCTCAATTTTATGAAGAACTGATCA 920
1021 ttaaatltaacaggagatcttaacagagagagccatagacatcccaagatattgttga 1080
921 TTTAAATTATAGGATTTCTAACAGAGAGGCGCATTAAGATTTCCAAAGATATATTGTAGA 980
1081 tgaagaagaatgtagagagcagagatgaaatgaagaagaagaagagagaaatacca 1140
981 CGAGGAAGATGTGAGACACAGATGAATATGAAGACAAAGAGAGAGAGAGATACCA 1040
1141 ggcacgttacagaagtgatccaaatctggcccggtatcccgtaaaagccaacctacga 1200
1041 GGCAGGTACAGAAAGTGTATCCAAATTTGCCCGGTATCCGGTAAAGCCACAACTATGA 1100
1201 agaacaaatgcgcatacagctgaggtgtlccaggtcagacatgtagagagacagtgta 1260
1101 AGACCAATGCGGATCCACGCTGAAGTCCGGGACACACAGAGAGAGCAGCATGA 1160
1261 gtttctcttgcaaacgctgagactagaagatccagatltctctgtctaagatgtag 1320
1161 GTTTCTTTGGCAATGCTGAATCGAAGATTCAGAGATTTCTGCTAAGGATGATAG 1220
1321 accatcaagcaaaagatctgtatctgaacgtlagagctgcaatggaanaaccacagatcgt 1380
1221 ACCATCAAGGCAAAAGATCTGATCTGAACGTAGAGCTGCAAATGGAAMCCAGCATCTGA 1280
1381 tttcaatgagaagaacttgagaaggtcagaagagaagaatttcaaccaagaagccctcaaa 1440
1281 TTCAATGGAAGAAACTCGAAGAGCTCAGGACAAAGTTTTCATCCAAAGAGCACCAAA 1340
1441 tcatagctctcccaaccctcgcgcgagagccctataccgcttgaatagaccagatgagcg 1500
1341 TCATGTCCTCTCTACCCCTCGCAGAGGCCCTATACCGCTCGATAGACCAAGCTGAGGCG 1400
1501 cgctgactccctcaggaanaacagcaccatctagatcccaactcgtcgttgaagaaacgaa 1560
1401 TGCCGACTCCCTACCGGAAACAACACCACTTAATCCCACTCTGCTGTATGGAAGAAAGAA 1460
1561 gcgagaaaaatggaacacatgttgaagatgtatcttgaagttcagagcagtcgagtc 1620
1461 GCGAGAAAAATGGAACAAATGTTGAGGATGATTTTGTAGTTACAGACCAAGCTGAGTGC 1520
1621 agtgaagccgcccccaacaaagccctcataatccaaagaagaagttaaatgcgcaggt 1680
1521 AGTGAAGCGCGCCCAACCAAGGCTCATTAATCCAGAAAGAGGTAAATGGCCAGGT 1580
1681 ttcactgagagctcggagagagagcttggtcatccacacctgaatatagaagctgttga 1740
1581 TTCACTGAGCAGCTCAGAGAGAGAGTTGGCATTCACGCTGAGATATACAGCTGTGATGA 1640
1741 tgtgagctggaagcgagagtgtagtgtagaagaggagacgtcaaaagggaaagaanaa 1800
1641 CCTGAGATGTAAGAGGAGAGAGCTAGTGAAGAAAGAGACAGTCAAAAGGAAAGAA 1700
1801 aactagtagcagggaggttltgcgattctaacaccaaggtc-gagagagaacaaagaag 1860
1701 AACTAGTAGCAGCGAGTTTGTGGACTCTAACACCAAGTCTGAGAGACAAAGAAAGT 1760
1861 gatgtactatggtggccactctlttgaaagaggaatttggaatgtctcgaagctcagataa 1920
1761 GATGTACTCTGGGGCCACTCTTTGACAGAGATTGGAAATGCTGACGCTCAGACTAA 1820
1921 ggaacttgaggtagatacctgtagttagtcaacaacccctaaacgaagagatagccaatgta 1980
1821 GGACTTGGGGTAGACACCTGTAGTACACACACTTAACGAGAGAGCGTACCTATGTGA 1880
1981 taagcaacctgtgaacctgagccatccaaagaatgagatcgctcaatttggctatctt 2040
1881 TGAACACCTGTGACCTGGCAGCATTCAAAGATGAGATGCGCTAATTTGGTCTATTTT 1940
2041 attaaataagcggtttaaagaatggaggtgtaacctcgagatcagaagcaatgctgggctt 2100
|||||

Db 1941 ATTAAATGAAGCCGTTTAAAGATGGAGTGTACCTCGAATTCAGAGCAATGCTGGCCTT 2000
QY 2101 aaaggtgtagagaggaagaatgactgaatcaggtcgacttctgcatcttattaccaaagt 2160
Db 2001 AAAAGTTGAGGAGGAAAGATGACTGATCAGTGCAGCTTGCTGCAATTTATTTACCAAGT 2060
QY 2161 aaaaaaagaagtttagctatactgtagacatcttagaccaggtgtagagttgga 2220
Db 2061 GAAAAAGAGAGTTTAGCTGATCTGTAAGACATCTTAGACACAGTGTATGAAGCTTGGA 2120
QY 2221 atggaatggagagctatcgcaagagccacatttgaggaagttacaacatctataga 2280
Db 2121 ATGGAATGGAAGGCTTATTCGAAGAGCCACATTTGAGGAAGTTTACAATTTATTTCTAGA 2180
QY 2281 atccaaacctgaaccacaagtttagctgtgtttcaaggccaattgagatctccag 2340
Db 2181 ATCCAAAGCCCTGAACCAACAAATTTAGCTTGTGTTCAAGGCCGATGGAGATGAGTCCTAG 2240
QY 2341 aataacctgaatgacgcatgcacaaactggaatccagttcttagcatttgatctcaaa 2400
Db 2241 AATACCTGATAGACATCATACACAGATCGAATCCAGTTCTAGCTCATTTGATTCGAAAA 2300
QY 2401 aatgagcgcctcctctatatacgcgttaaccaccaagaagtcgcatgctgaaggaatgt 2460
Db 2301 AATGGAATCGCTCCGCTATATCGGTTACCTCCCATGAGTCTGCGATGCTGAGGAGATGT 2360
QY 2461 ccggagctctctactcggagaagcttcaataaactagtgtttgacaaggttgctacca 2520
Db 2361 CCACAGCTCTTATCTGAGACAGCTTCAATTAACATGATGTTTACAGAGTTGGTCACCA 2420
QY 2521 gttagatgattacaattttggaagcaagaatcctccctccaggaagaatggagagccag 2580
Db 2421 GTTAATAGTTACATTTTGGAGCAAAAGATCTCCCTTCAGGAGAAATGGAGGCCAG 2480
QY 2581 gaatccttatgtagaattactcctccagabagaagtgtataaaataagagaagaac 2640
Db 2481 GAATCCTTATGTTAAATTTACTTCTCCAGACGAGATGATAAAAACAAGAGAAAGAAC 2540
QY 2641 aaaaaagtcagaaaaaactttggaaccacaatggaaccagaacttctattctccgt 2700
Db 2541 AAAAAGTCGAAGAAAACCTTTGGAACCCAAATGGAACAGACTTTCATTTATCTCCGT 2600
QY 2701 ccacggaagaagaatccgttgaacgaatgctggaatcaatcccttgggaatcagatagat 2760
Db 2601 TCCACCAAGAGATTCGGGAGACGAATGCTCGAAATCAACCTTTGGATCAAGCTCGAGT 2660
QY 2761 tcgagaagaagaagaatccttagagagaatttaattgaatttgaacagcttgct 2820
Db 2661 TCGAGAGAGAAAGATGTAATCTAGAGAGATTTTAATTGAATTGMAACAGCTCTGTT 2720
QY 2821 aggtatgtagccgactggtatagctgcagagaccatgatatgtctccatctgcaactccc 2880
Db 2721 AGATGATGAACCAACACTGGTACAAACTTCAGACCATGATGCTCTCATTTGCCACTCC 2780
QY 2881 tcgaccttcccacatctgcgccggagagcagctccatcggaagagagcccaagcgagct 2940
Db 2781 TCCACCTTCTCCATATATGCGCAGGAGACAGCTGATGGAAGAGCCCAACACGAGGTT 2840
QY 2941 gcaaaagtcgaagaagaatgaatgacagtgaaagtgtctgactcagactgcagagatgagct 3000
Db 2841 GCAAAAGCTCGAAGAGATATGTAACACGAAATTCGCAATATGACTGCGAGAGCGCGT 2900
QY 3001 gggagtagtgtagattatagacacaaatgscgagatcttaaaagctcaagttgtggt 3060
Db 2901 GGGAGTGTGTGATTAACGACACGATGGCCGTGATCTTAAAGCTCCACATTAATACGT 2960
QY 3061 gccagaacaagtatgtcatcaaatcatctgtcaccaatcaggtctccctcatcgagtaga 3120
Db 2961 GCCAGAACAAAGTATGTCATCAAAACCATTTGCTACATCAAGGCTCTCTCATTCAGAGTAGA 3020
QY 3121 tgttatagaagaacaaggtcatatgctgctagtgtgcccctctctcctaagaagaatgtga 3180
Db 3021 TGTATTAGGAAGGACTAGGTGATGCTGCTAGTGTCCCTCTCTCTCAAAAGGAATGTGGA 3080

QY 3181 acaggggcaaccgagggacagtgctactggccattacaacaatltagccgaatgtag 3240
Db 3081 ACAGGGGCTTCGAGGGACACAGTGTCTACTGTGSCATTTCATAATACATATAGCCGATGGACAG 3140
QY 3241 acacgctgcatatgtagtgaacacactactctcttagaagagc-----aggatgtgaagc 3280
Db 3141 ACACCGTGTATGAGACGACCATTTACTCTTCAGAGAGACAGCATTTTCTCACTTAC 3200
QY 3281 -----aggatgtgaagc 3294
Db 3201 TCGCTTCGACACAGCAGACAGCATGAGCATCACACAGGATGGAAGGATTTGTGAAGC 3260
QY 3295 agcagaatagaacaacatatacagatccagatcaacaagaacaagcgtctcccaagagc 3354
Db 3261 AGCAGATGACAGCCATATTCAGATTCAGATCAATCAACGAACAGGCCCTCTCTAGAGCG 3320
QY 3355 gacacacaccgcctccagatcctctgaacgttccctgatacaaaacctcatgaagtcgatgac 3414
Db 3321 GACACACACCCGCTCCAGATCTCTGAAAGCTGTGATACAAACCTCATGAGGTGATGCC 3380
QY 3415 ttaataatgactggaagaatctgcccctccttaacctgtactatcgaggtctcaacctcg 3474
Db 3381 TTTCAATTAATGACTGGAAGATCTGCCCTCTTCACTGCTTATTCGAGGTCTCACCCCTCG 3440
QY 3475 taccgggtctgtccagacaagcccatcaagttactcgcgggacaagaagagggccgaca 3534
Db 3441 CACTGCTCTGTCCAGCCAGTCCGTCAATCTACTCCGTACAGAGCAAGGCGCGCA 3500
QY 3535 gcttccacagcttccacaaaggaacattggaagagaggtgctatggaatalagagagag 3594
Db 3501 GCTTCTCAGCTCCCAACAAAGGAACACTGGAAG----- 3536
QY 3595 aaatgcacaatgaaacttaacaatacaaaacaggtagccgatalcagaacccagactgga 3654
Db 3537 ----- 3536
QY 3655 gaaagattaccattcgaaatgcgtcagatagggatccacatagagggagatactgt 3714
Db 3537 ----- 3536
QY 3715 ttccactaaatccctcgcgacgtgagtgaatgtagtactgtcgcgtttcaagactagtag 3774
Db 3537 ----- 3536
QY 3775 tgcctcgtttcagaagacaagaactacatgctccatcaagaagcgccgagaagaaa 3834
Db 3537 ----- 3536
QY 3835 caggaaatatgctgtcttaccatccaaatgcaaaacagacagatlggcygtcgggaa 3894
Db 3537 ----- 3536
QY 3895 gaacttgacaagaagacagacatcagtggaagacatgtgtcacttggagaagaaatgacgy 3954
Db 3537 ----- 3536
QY 3955 cagcagctccagacatcgagtgaggcgccctgggtacacatgagcaagaagcgatctag 4014
Db 3537 ----- 3536
QY 4015 catltygggccaanaatgtagctatgtgtgtctcctcaaggaagatcgagtgctctca 4074
Db 3537 ----- 3536
QY 4075 actcagccaacacgaaggaaggaatlaanaagctacggaagcactgttcagagaagaacagga 4134
Db 3537 -----AGAGAGAGGTAAAGGCTTACGAGCATGCTTCCAGAGAACAGACGGA 3581
QY 4135 gaccgggtcagcagtgtagatgtaggaagcgtgatagaaccgccaagccagccggaatccac 4194
Db 3582 GACCGGGCTAAGCAGATGAGATGAGAACTGTGATGACAAGAGCGCCAGCCGGAAATCCAC 3641

Db 228 -----AAAGCTGCATCAACAAATTTGAAATGATAAGGA 260
QY 361 gcaagtcaagaagatgaggaggaatcgcagcagcagcaagagcagaaggtgtgcccc 420
Db 261 GCAGATGCAAGAAATGTGGAGAGAAATCACAAACAGCAGACAGAGCGCGCGCC 320
QY 421 gacccgtgcatctgcacaagaacaatttgcagatgagatgaggccataattgttcta 480
Db 321 GACCTGCGGCATCTGCCACAGACAAATTTGCGATGATGCGCGCCATACCTTTATA 380
QY 481 ttgcacaaccaagttctgctgcagatgltggaagltcgagltgcttaccgtccaacaggt 540
Db 381 TTGCCAAACCAAGTTCTGCTGCTGTGTGGAGGTGAGTGTCAATTACGTCAACCAAGT 440
QY 541 tatgtgtgtgttaatttctgcccgaacaacaagaattccctaactaatacagagcatg 600
Db 441 TATGTGGGTGTAAATTTGTGCGAAACCAACAAGAAATCTCACTAAGTGTGGCGGTG 500
QY 601 gtttataatagtggtctaaacactgcaaacctgatacaaggltcctcgagggt 660
Db 501 GTTTTACAAATAGTGGGTCTAATACACCGCAGCCTGATCAAAAGGCTTTCCAGGGCT 560
QY 661 tcgaaatgagaagccctcaggaagaagaagaactacagagcagcccaagttcca 720
Db 561 TCGCAGTGAAGAAAGCCCTCAGGAGAAAGCAAACTCATGAGCAGACGAGTTCCA 620
QY 721 aggaagcccaagttgactatcagttaccgtcagttgagaagagccagatcaggtcac 780
Db 621 GGGACCCCGGTGATCTCACTACCTGCAATGAGAGGTGAGAGCTCATGTGGCTCAC 680
QY 781 aagacagagttactataaataatgaltcagaagtgaaagcaccagatcgcagttgacg 840
Db 681 AAGACAGATTTCTATTAAAGATGCTCAGGAATGAAACACCAAGTCCCACTGACATGCC 740
QY 841 ttcagacagaanaacgaagttccatcagltgltcagaggtacaaatcgaaatcagaagaag 900
Db 741 TTCAGCAGAAAAAAGAGTCCGTCAGTGTCCAGGATCAAAATGGAATGAGACCAAG 800
QY 901 tgaagaagaagaggtactcacaagatagtcttccctcagatggtacaaatgcgaatctcc 960
Db 801 TGAAGAAAGAGGAATATTACAGATATGTTCTTCAGATAGCAAAATGCTTAATCTCC 860
QY 961 ttcagatctgctcagatcagatcagcgttgaagcccaatttataagaacctgtgca 1020
Db 861 ATCAGATTTGCTGATAGACGATCTCAACGTGAACCTCAATTTTATGAAGAACCTGATCA 920
QY 1021 ttttaaatcagaggtacttaacagagaaggccatagacattccaaagatlatltgga 1080
Db 921 TTTAAATTTAGGATTTCTAACAGAGAGGCCATAGACATTTCCAAAGATATTGTAGA 980
QY 1081 tgaatgaagatgtgagagcagaagatgaataatgaagaagaagaagaagaataacca 1140
Db 981 CCAACGAAGATGTGAGAGCAGAGATGATATGAAGCAAAAGAGAGAGGAGGATACA 1040
QY 1141 ggcacgctcagagaagtgatcccaaatctgcccgttaccctgaagccacaacctacga 1200
Db 1041 GGCACGCTCAGAGAGTGAATCCAAATTTGGCCCGTATCCGTTAAAGCCACAACCTTAGA 1100
QY 1201 agaaacaatgcatccacagctltgaggtgtgccagagcagacatgagaagaagcaggtga 1260
Db 1101 AAGAGCAAAATGCGGATCCAGCGTGAAGTGTCCCGGGCAGCAGAGAGAAGAGCAGATGA 1160
QY 1261 tgttctcttggcaaacgctggaactagaagaatccagaattctctgtaagaatgagtag 1320
Db 1161 TGTTCCTTTGGCAAAATGCTGAACATAGAGATTCACAGATTTCTCTGAAGAGTGAAG 1220
QY 1321 accatcaagaagcaagaatctgtatctgaacgtagagctgcaatgtaaaaaaccaagatcta 1380
Db 1221 ACCATCAAGGCAAAATCTGTATCTGACGTAGAGCTGCAATGGAACCAAGCGATCTGA 1280
QY 1381 ttcacatggaagaagcctcagagaggtcctcagagacaagttctatccaaagaagactcaa 1440
|||||

Db 1281 TTCAATGGAAGAAACTCGAAGAGCTCAGGACAAAGTTCTTATCCAAAGGACACACAA 1340
QY 1441 tcatagtcctccacccctcggcgagccclataccgcttgatagaccacaatagagcg 1500
Db 1341 TCATAGTCTCTTACCCCTCGCAGAGCCCTATACCGCTGATAGACCAAGCTGAGGCG 1400
QY 1501 cgtgactccctacggaacaagcaccatagatcccaagctctgctgtgagaagaacga 1560
Db 1401 TGCCGACTCCCTACGGAACCAACACACACACTAGATCCAGCTGTGCTTAAGGAACGAA 1460
QY 1561 gcgagaaaaaatggaacaatgltlaaggaatgactcttgatltcagaaccaagtcagtc 1620
Db 1461 GCGAGAAAAATGAAACAATGTTGAGGAATGATTTCTTGTAGTCAACCAAGTCTGAGTC 1520
QY 1621 agtgaagccgcccccaagagctcataatccaaagaagaggtgaatagtcgccaagt 1680
Db 1521 AGTGAAGCCGCCCCCAAGGCTCATTAATCAAGAAAGAGATGAATATGCCCCAGGT 1580
QY 1681 ttcactgaagcagctcggaagagagctggtatccacacccgtgaatatagaagctgtga 1740
Db 1581 TTCACGTGACGACGCTCAGAGAGAGAGTGGCATTCACGCCCTGAGATATCAAGCTGTATGA 1640
QY 1741 tgtgagcttgaaagcagagatgtgagtgagaagaagggacagttcaaaaaggaagaa 1800
Db 1641 CGTGAGATTGAAGCGAGAGCGTAGGTGAGAAAGGAGACA----- 1681
QY 1801 aactagtgagcgagaggttctgcatcttaacacacaggtctgagagacaaaagaag 1860
Db 1682 ----- 1681
QY 1861 gatgtactatgtygacactcttgaagaagattgaaatgltcagccctcagatlaa 1920
Db 1682 -----TGAGTACAGCTGTGTGAGACATGCGTCTTG 1712
QY 1921 ggaacttgaggtagatacctgtgtagcacaacctaaacgaagagcatagccaatgta 1980
Db 1713 G-----CATAGCAGTGAAGCCATCCCAATGTC 1739
QY 1981 taagacacctgagacctgagccatccaaagaatgtagatcgcgctaatttgctatatt 2040
Db 1740 TTTGACCCCTGTGACCTGTGAGCCATCCAAAGATGAGATGCGCTAATGTGCTATATT 1799
QY 2041 attaaataagcgttlaaagaatggaatgtagtctcagacatcaagagcaatgctcgagct 2100
Db 1800 ATTAAATTAAGCGTTTAAAGATGGAGGTGATCCCTCGAGATTCAGAGACAAATGCGGCTT 1859
QY 2101 aaaggtgtgagaagaagaatgactgaatcgaatcgaacttgtgcatlttataccaaagt 2160
Db 1860 AAAGGTTGTAGGAGAAAGATGACTGAATCAGGTGCACTTTGTGCATTTTATCCAAAGT 1919
QY 2161 aaaaaaagaagtttagcttgatctgtagacatcttaagccaggtgtagaagcttga 2220
Db 1920 GAAAAAAGAGAGTTTGTGCTATCTGTAGGACATCTTAGACCAAGTGAAGCTTTGGA 1979
QY 2221 atggaatggaagctatctgcaagagcgacacatttgagaagtttacaacatlatctaga 2280
Db 1980 ATGGAATGGAAGGCTATTGCAAGAGGCGACATTTGAGAGAGTTTACACATATTCTAGA 2039
QY 2281 atccaaacctgaaaccaagltgagcttgttltcaagccaatttggagataltccag 2340
Db 2040 ATCAAGACCTGAACCAACAGTTGAGCTTGTGTTCAAGGCCGATGAGATATGAGCTAG 2099
QY 2341 aataactgtatagacgcatgacaaactggaatccagttctagctcatttgaatccaaaa 2400
Db 2100 AATACCTGTATAGCAATGCAACATGGAATCCAGTTGTAGCTCAATTTGAATCTCAAA 2159
QY 2401 aatgacccgtccctctatctcgttlaactcaccatgaagtcctggaatgctgcaagagat 2460
Db 2160 AATGATTCGTCCTGTATATCGGTATACCTCTCCATGATGCTGTGCGATGCTGAGAGATGT 2219
QY 2461 ccgcagttcttacttgagacagttccaataaactatglttgacaaggttgtgtacca 2520
Db 2220 CCACAGTTCTTATCTGACAGCTTCAATAAACATATGTTTGAACAAGTTGGTCAACA 2279
|||||

QY	2521	gttatagttacatttttggagcaaaagatctcccttcagaggaagatggagagcaag	2580
Db	2280	gTTAAATAGTTMCATTTTGGGAGCAAAAGATCTCCCTTCAGGGAAAGATGGAGGCCAAG	2339
QY	2581	gaatccthaigttaagatttaactcttcctccagatagaagtataaaataagagaagac	2640
Db	2340	GAATCCTTAATTTAAAAATTACTTCCTTCAGACAGAAAGTATAAAAACAAGAGAGAAC	2399
QY	2641	aaaacagtcagaanaactttcgagaccacaatgagacagacttcatltacttcctgt	2700
Db	2400	AAAAACACTCAAGAAANCTTTGGAAACCCAAATMGAAACGAACTTCATTATTCCTGCT	2459
QY	2701	ccaccgaagagaattcsgltgacgaatgctggaatatccctttgggtcaagctcaggt	2760
Db	2460	TCACCGAAGAGATATCCGGGAACGAATCTCGAAATCAACCTTTGGATTCAGAGCTGAGT	2519
QY	2761	tcgagaagaagaagacggaattccttagagagatacttaattgaattgagaacagctttgct	2820
Db	2520	TCGAGAGGAGAAATGTAATTCCTCTAGAGAGATTTTAATGAAATTGGAAACGCTCTGTT	2579
QY	2821	agatgatagccgcacgtgltataagctgcgaagcccatgatatgtctcctcatlgtccactcc	2880
Db	2580	ACATGATGAAACACACTGGTTACAAATTCAGACCCATGATGTCTCTCATTTGCCACTCCC	2639
QY	2881	tcgctctcccatctctbccccggagagagcttcatagtagagagcccaagcagagct	2940
Db	2640	TCACCTCTTCATATATGTCACCGAGNACGCTGCATGAGAGAGCCCAACGCGAGGTT	2699
QY	2941	gcaaaagtcgaagaataatagtgacaagtgaagtgltctgactacgactgcgagatgagct	3000
Db	2700	GCAAAAGCTCGAAGAGAAATAGTGACAGCGCAATATCCGATATCGATGTGAGAGAGCGGCT	2759
QY	3001	ggagtagtgatgaatatatcgacaaatgcccggagatctcaagctcagcttgcgt	3060
Db	2760	GGGAGTCTGTGAGATTACCGACAGCATGCGGTGATTTCAAAGCTCCACATTATCACT	2819
QY	3061	gccagaaacagttatgtcatcaaatatttgtcatccatcaatgaagttcctcatcagtaga	3120
Db	2820	GCCAGAACAAAGTATGTCAATCAAAACCATTTGCTCACCATCAGGCTCTCCATGAGATAGA	2879
QY	3121	tgattatagaagaagaagtcatgtgtcgcgttagtgcccttcctccctcaaggaatttga	3180
Db	2880	TGTTAATAGGAAGGACATGAGTCATGAGTGCGCTATGTGTCCTCTCTCCANAGAAATGTGA	2939
QY	3181	acaggggcaeccgaggagacaagtctactgtgccaattcaacacaaattagccgaatggatag	3240
Db	2940	ACAGGGGCTTCGAGGGAGACAGCTGCTACTGGCCATTTCAAATATACCAATTAGCCGATGAGACAG	2999
QY	3241	acaacgltgcatgagatgacactactcttcatagatagagaaggaattgtgaacgagcag	3300
Db	3000	ACACCGGTGTCATGGAGACACATTACTTCTTAGAGAGAGACAGGATGTGTAACGAGACGA	3059
QY	3301	tagaagacccatcaacagatcccatcaacagacaacaaacggcctcctctagagcgagcac	3360
Db	3060	TAGACAGCCATATCAGATCCAGATCAACGAACAAACGAGCGCTCTCTTAGAGCGGAGCAC	3119
QY	3361	caaccgctccagatcctctgaaagctctgtatacaaacctataagtgctgagttcatt	3420
Db	3120	CACCCGCTCCAGATCTCTGACACCTGTGATACAAACCTATGAGGTGATGCTTCATT	3179
QY	3421	aatgactggaagatctbccccctcttcaactgctctltagagttcaactcgttaccg	3480
Db	3180	AATGACTGGAAAGATCTGCCCTCTTACTGCTTATTCAGAGTCTCACCTCGCACTGG	3239
QY	3481	gtctgtccagacaagcccatcaagtlactccggaagacagcagcaaggggcgcgaagcttc	3540
Db	3240	CTGTGTCAGACCGTCCGTCACAACTACTCCCGTAAACGAGAGGAGGCGCGCAGCTTC	3299
QY	3541	acagcttccacaaagggaacatctggagagaaagtgctatgata tagagaggaatctg	3600
Db	3300	TCAGCTCCCAACCAAGGAAACACTGAGAAAG-----	3329

Qy	3601	coaaatgaactlaacaaatacaaacagtgacccggtacagaccocagactcgtgacgaaga	3660
Db	3330	-----	3329
Qy	3661	ttaccattcgaagtatcgctcaagatggtgatccacatagaggggacagatactgtttccac	3720
Db	3330	-----	3329
Qy	3721	taaalccctcgacagtgatgaatgatatctgcggttccaaagactagtatgtctc	3780
Db	3330	-----	3329
Qy	3781	tcgttcagcagcaacaagctaaatgtccgtccaatcagagcgcgcgagaggaacaagaa	3840
Db	3330	-----	3329
Qy	3841	aatcagtgctttacatccaaatgcaaaacagacagatbggctgtcgtgggaagaactt	3900
Db	3330	-----	3329
Qy	3901	gaccaaagaccacgacatcagtgtagagacaatgtgctccactggaagaagatgacgcagcca	3960
Db	3330	-----	3329
Qy	3961	gtccgacactgcagtgtagccctggtgtacaaatggtgcaagaagcgcgcgtatagcatlgtg	4020
Db	3330	-----	3329
Qy	4021	ggccaatatgttagctatgttgtgtctctccacggaanaatgcgagtgtcctccaaactcag	4080
Db	3330	-----	3329
Qy	4081	ccaaaccgaagggagggatlaaaaagctacggaagcactgtgtcagagaagacagagaccgg	4140
Db	3330	-----AGGAGGAGGTTAAAGCTACGCACACTGTCTCAGAGAAGCAGGACCGG	3380
Qy	4141	gtctagcagtgtagagatgagaaactgtgatccgcgcgcagccacgcgcggaatccacagatg	4200
Db	3381	GCTAGCAGTGGAGATGAGGAAGTGGATGACAMACAGCGCCAGCGGGAATCCACGGATGG	3440
Qy	4201	caagatgaacaagctatagctcgcgaagaagaatctgatatcttcctcgtgggtccgcctgtgcctc	4260
Db	3441	CAGATGAACAAGTTACAGCTCACAAGAAATCTGATCTTTCCTGGGGTCCGCTGGCCTTC	3500
Qy	4261	tgaacagcagttcagtgattctctcgtgatatgcctggtgcgcctgtcgaagctagtgtgagccca	4320
Db	3501	TGACAGCCAGTTGAGCATTTCTTGATGGTCTCGGCCCTGCTCAGCTTAAGTGGACGCCA	3560
Qy	4321	gaacctgtgcactccctcacaatggtgatgacatctcaggtgtggaatgatgataaaaaggaca	4380
Db	3561	GACCTGCGCACTCTCCAAATGGGTGACATTCAAGTGGGATGATGATMAAGAAAGGACA	3620
Qy	4381	gcttgaggttagaanaatccggtgcgcgcgcgcgcctgtgtgttaaaccaaggttccaaagacct	4440
Db	3621	GCTGGAGGTAGAGATCATCCGGGCTCGAGGCTTGTGTAATAACACAGGTTCCAAGACACT	3680
Qy	4441	gccgcagccagtatctcaagtggtatctgtttagaaaaagagagtgatcgtacacaaanaagaa	4500
Db	3681	GCCAGCACCATATGTCAAGGTGTACTCTGTTTAGACAAGGAGTCTGCAATAGCCAAAAAGAA	3740
Qy	4501	aaccaaggtgtgcgaagaagaccctcgtgagccctgttaccagcagctctgttcctcgagga	4560
Db	3741	AACCAAGTGGCGAANAAGACGCTGGAGCCCTGTATCAGACGCTGTTATCTTTCGAGA	3800
Qy	4561	gaagcccccagggaggtgtgttaacagatcatgtgtctggggaatatatgtgtatgataca	4620
Db	3801	GAGCCCCAGGGGAAGGTGTACAGATCATTTCTCTGGGGAATTAATGCCCCGCAATGATCA	3860
Qy	4621	caaatccttatgtgagatggtccacgaatactttagatgaatgtgaaactatccaaatggtt	4680
Db	3861	CAAACTCTTTATGGAGTGGGCCCAAGATACCTTTAGATGAACCTGGAACATATCAACATGTT	3920
Qy	4681	gattgatatgttcaaacctctcctcctctcctcctgtagatccaaactcgtgacacctt	4740

|||||
Db 3921 AATTGGATGTTCAAAACCTTCCCTCCCTCCCTAGTAGATCCAAACCTGGCACCTCT 3980
4741 gacaaagaagatccccaatgctctctggaagaagttctaccgaactcttactctgctc 4800
3981 GACAAAGAAGACCTTCCCATCTCTCTGTGAAAGTTCTACTGACCTTCTACTCTGTTTC 4040
4801 atagaacataaactctgtcacacaacagcagatatacaaaacagaagaacagca 4860
4041 ATGCACTATTAACCTGTTGTACACACACACAGCCTTACAAAACAGAGAACACCACC 4100
4861 caagctgaagacccctgtgaacacatgcatgctgtagtctgtctacagaagccacagctc 4920
4101 CAGGTGG-GGAACCTGGTCCCACTGCATGCTGATGTTGTGTCTCGGAGCCACGCTCTA 4159
4921 gggataccagaagcagtcgtctgtctcagaagaag----tcgtaacatctgtgcctagca 4976
4160 GGGATTAAGAGCCATCTGTGCTCTACAGAGAAGCCATCCGTGCACAGTGTGCCCTAGCA 4219
4977 aagg 4980
4220 AAGG 4223
RESULT 7
AF199323 5031 bp mRNA ROD 04-JUL-2000
LOCUS AF199323 Rattus norvegicus RIM2-2A (Rim2-2a) mRNA, complete cds.
DEFINITION AF199323
VERSION AF199323.1 GI:8925859
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 5031)
Wang Y., Sugita, S. and Sudhof, T.C.
The RIM/NIM Family of neuronal C2 domain proteins. INTERACTIONS
WITH TAB3 AND A NEW CLASS OF SRC HOMOLOG 3 DOMAIN PROTEINS
J. Biol. Chem. 275 (26), 20033-20044 (2000)
2 (bases 1 to 5031)
Wang Y. and Sudhof, T.C.
Direct Submission
Submitted (27-Oct-1999) Center for Basic Neuroscience, The
University of Texas Southwestern Medical Center, 6000 Harry Hines
Blvd., Dallas, TX 75235-9111, USA
FEATURES
source location/Qualifiers
1. 5031
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1. 5031
/gene="Rim2-2A"
52. 4110
/gene="Rim2-2A"
/codon_start=1
/product="Rim2-2A"
/protein_id="AAF81645.1"
/translation="MSAPLGRPGRPAPPAASOPPPPEMPDLSHLTTEERKLTIOVM
DROKKEERKBOVSLKILHOEFMYKEQVKMGEEFSGOEOEOADATCGICHTKPAD
GCCNGCSTCOTFCARCGRVSLRKNKVMVNCNLKROEILITKSAWPTNNGSNTPD
QPDKALGLRSEEPQEKAKLHEOTPOGPPGDSVAVERGRAHGLTRDSING
SGMHHQIASDMPDRKRSPSVSRDNRKRYDQSEEREYSQYVPSDTPMRSPDYADR
RSQREPOFEEDPHLNTYDNRNKRHRSKEYIVDEEDVSRDEYEQREEREEOAARY
SDPLARYRPOPEEOMRIHAEYRARHERHSVSLANLELDSRLSLCHMDPS
ROVSEPRRAMENORSYSMERTEAQOSSYPORTTNSPPTPRSPSPIDLPDLR
ADSLRKQHHLDPSAVKTKRKMETMLRNDLSLSQSSVPRPPRPKSKGGMK
QVSLSSSEELASPEYTSQDVEITESVSGKDEYSLWEHAKMSSSPASLHP
VTWOPSKDGLRIGRLILNKLKDGSPADSGAMLGKLVGKMTSGRLCAITVK
KGSITADYVGHRLPGDEVLEWNGRLQAGATFEFVYNIIESKEPEVELVSRIGMP
RIPDSTHAQLSSSSSSSFESQKMDRPSISVTSFMSPGMLRDVQFLSGQLSIKLPFKV

GHOLIIVTILGAKDLPSREDGRPNRYKIYELPDRSDKNKRRTKTVKLTLEPKMNOTE
IYSPVHRPEERMELEITLMPDQARRESESEIGLEILEFALLDEDEHMYKLOHD
VSSLPPIRSPRYMRRLHGESPMRLRQHSRIRISDSYDPCDEGVYSPYRHDR
DLOSTTISVPEQVWSSNHCSPSGSEHPRDVIGRTSMBSVPPPORNVQIGRTRAT
GHYNTISRMDRHYWDHYSSERDSHFLTLPPSRHRYSEHHNRGDRDCEADRDQYH
RSTEDRPLERTTTRSSERADNTNMRSPSLMTGSAAPSPALSHSRPTGSVO
TSPSTPYTGRRGROLPOLPKGTLEREGGKLRSTVQSTGLAVEKRMNTROAS
RESTDGSMSNSYSSGNLIFPGVRLASDSQSFELDGLPAQVGRQTLATPAMGDIOV
GMMDKQOLPEYITIRARLVYKPGSKTLPAPYKLYLDNQCIAKKTKVARKTLEP
LYOOLSFEEESPQGVLOIYVGDYGRDHSFMKVQAQILLDELELSNVIQWFKLEP
PSSLVDPLTLPTRASQSSLESSTPSYRS"
BASE COUNT 1503 a 1166 c 1246 g 1116 t
ORIGIN
Query Match 57.4%; Score 2860.4; DB 94; Length 5031;
Best Local Similarity 78.0%; Pred. No. 0;
Matches 3938; Conservative 0; Mismatches 261; Indels 851; Gaps 7;
1 gcttcctagggtggtgctgcgcgcgcaacatgctgctcgtgggcccggggcgg 60
21 gcttccctTAGGGTGGTGGCTCCACCAACATGTGGCTCCACTGGGGCCCCGGGCGG 80
61 ccggctcccaaccggcgagcctctcaacctctccgcagcccgagatgcgagcctcag 120
81 CCCGGCTCCACCCCGGGGGCCTCTAGCCTCTCCGCGACCCGAGATGCCGACCTCAG 140
121 ccaactcaaggaagaagagaagaatacctcgtgctgtcatgagtgcgtcagaagaaga 180
141 CCACCTCACGGAAGAGAGAGAGAAATCATCCAGCCTGCATGATGATGTCAGAAAGA 200
181 agaggaagaagagcagtcgctgctcaagaatcaagaagaacaagaacacaccgacaca 240
201 AGAGAGAAAG 227
241 gtgttcctcttagtggatcactgactgtaactgtaactgtaactgtaactgtaactg 300
228 ----- 227
301 acaaccaatgagaagagagcccaagaacaagtgcgacccaacaattgaaatgataaga 360
228 ----- -AAAGCTGCATCAACATTTGAATGTATAAGA 260
361 gcaagtcagaagaatgaggaagaagaatcgacagcagcagaagaagcagaaggtgatgccc 420
261 GCAAGTCAAGAAATGGGAGAGAAATGCACACAGCAGCAAGCAAGCAAGGCGCAGCC 320
421 gacctgtgcatctgcacaagaacaatttcagatgagatgagtgccatgcttca 480
321 GACCTGCGGCATCTGCCACACACAAATTTGCAGATGGATGGCGCATTAAGTTCATA 380
481 ttgccaacaacagttctgtctcgatgtagtgagtgtagtgcttcaactcaacaagaat 540
381 TTGCCAAACCAAGTTCGTGCTGTGTGGAGGTGCAAGTGTCTATAGCTCAAAAGGT 440
541 tatgtgggtgtgtaattgtgtgcgaaacaacaagaatccttactcaatcagaagatg 600
441 TATGTGGGTGTAAATTTGTGTGCCGAAACACCAAGAAATCTACTPAAGTGGCGGGTG 500
601 gtttataataatgagttcaacaacatgcaaacctgtatcaaaaagtctcctgaaggtc 660
501 GTTTACATATATAGGTGCTAATACACCGCAGACGCTGATCAAAAAGGCTTTTGAGGGCT 560
661 tcgaataatgaagaagccctcagagaagaagaacaactacacagcaagcccaagttcca 720
561 TCGAGTAGAGCAAGCCCTCAGAGAAAGCAAAACTGCATGACAGCAGCAGAGTTCGA 620
721 aggaagcccaagtgactatcagactcagtgactgagtgagaagccgagctatgagctcac 780
621 GGGACCCCGGTGACTCTCATGATCTCAGTTGAGAGAGGTGAGCTCATGAGGCTCAC 680
781 aagaagataataataaaatgatacagagatgaagcacaacatggcagttacatgccc 840

Db 681 AAGACGATCTATTAGAAATGGCTCAGGAATGAAGCACACAGATGCCAGTGACATGCC 740
 QY 841 ttccagacgaaacgaagccatcagtgctccagggatccaatccgaatcagcagcaag 900
 Db 741 TTCACACAAAAAAGAGTCGTCAGTGTCCAGGGATCAAAATCGAAGATACGACCAAG 800
 QY 901 tgaagaagaagagactacacagiatgttccttcagatggtatacctgccaaatctcc 960
 Db 801 TGAAGAAAGAGGAGATATTCACAGTATGTTCTTCAGATAGCACAAATGCTTAGATCTCC 860
 QY 961 ttcgagattatgctatagaagatctcagcgttgagccctcaatttataagaactgttca 1020
 Db 861 ATCAGATATGCTGATAGAGATCTCAAGCTGAACCTCAATTTATGAGAACTGTGATCA 920
 QY 1021 tttaaattcaaggatctctacaggaagagccatagacattccaaagatlatatgttga 1080
 Db 921 TTTAAATTTATAGGATTTCTAACAGNAGAGCCATAGACATCCAAAGAGTATATTGAGA 980
 QY 1081 tgaatgaagatgtgagagcagagatgaatatgaagacaaagagagagagagaaataga 1140
 Db 981 CGACGAAGATGTGAGAGCAGAGATGAATGAATGAAGAAAGAGAGAGAGAGATACCA 1040
 QY 1141 ggcagcctaaagaagatgataccaatctgcccgttatccgttaaagccacaacctaga 1200
 Db 1041 GGCACGCTACAGAAAGTATCCAAATTTGGCCGGTATCCGGTAAAGCCACACCTATGA 1100
 QY 1201 agacaacatgcgacacgctgagatgttccaggtgacagagacagagagagagagagatga 1260
 Db 1101 AGAGCAAAATCGGATTCACCGTGAAGTGTCCGGGACAGACAGAGAAAGCACAAGTGA 1160
 QY 1261 tgttctcttgcaaacgctgacatagaagatctcagagatctctgcctaaagatgatag 1320
 Db 1161 TGTTCTTTTGGCAAAATCTGTAAGTGAAGATTCAGGATTTCTGCTTAGAGATGATAG 1220
 QY 1321 accatcaagcacaagatctgtatctgtaacgttagagctgcaatgtaaaaccaaacyatga 1380
 Db 1221 ACCATCAAGCAAAAGATCTGTATCTGAACGTAGAGCTGCAATGGAACCAAGAGATCGTA 1280
 QY 1381 ttcaattgaaagaaactcgaagagctcagagagacaaagttctatccaagaagacctcaaa 1440
 Db 1281 TTCAATGGAAGAAACTCGAAGAGCTCAGGGACAAAGTTCTTATCCAAAGGCCACAAA 1340
 QY 1441 tcaatagttctccacccctctgagcagacataccagcttgatagaaccagacagagagcg 1500
 Db 1341 TCATAAGTCTCTTACCCCTCGAGAGACCCCTTATACCGCTCGATAGACAGAGCTGAGGCG 1400
 QY 1501 cgtcgaactccctacgaaacagcaccacttagatcccaagctcgtgtgtaggaagaaacgaa 1560
 Db 1401 TGCGGACTCCCTACGGAACAAACACCACTTAGATCCAGCTGTGCTGTAAAGGAAACGAA 1460
 QY 1561 gcgagaaaaaatgaaacacatgtttaagaaatgattcttgagttcagaccagtcgaagtc 1620
 Db 1461 GCGAGAAAAAATGGAACCAATGTGAGGATGATTTTGTGAGTTGACACAGTCTGAGTGC 1520
 QY 1621 agttagggccgcccccaacagagcctcataatccaagaagaagagatgaatgtgcagagt 1680
 Db 1521 AGTGGGCGCGCCCCCAACAGGCTCATTAATCCAGAAAGAGAGGTAAATGCGCCAGGT 1580
 QY 1681 ttcaattgacagctcggagagagagcttgacatccacacctgagtatatacaagctgtgtga 1740
 Db 1581 TTCACTTGAGCACCTCTAGAGAGAGAGTGTGGCATTCACGCTCGAATATCAAGCTGTGATGA 1640
 QY 1741 tgtgagactggaagagagagatgttgtagtgaagaaggagacagtcacaaaggaaaaagaa 1800
 Db 1641 CGTGAAGATTGAAAAGGAGAGCGTAGGTGAGAAAGGAGAA 1681
 QY 1801 aactagtgagcagggagttgttcgcgatatctaaacacaggtlctgagagacaaagaaag 1860
 Db 1682 ----- 1681
 QY 1861 gatgtactatgtgtgcccactcttggaagaagatttgaagatgtctgagcctcagattaa 1920
 Db 1682 -----TGAGTACAGCTGTGTTGAGAGCATGCGCTTTG 1712

QY 1921 ggaacttgaggtagatactctgtagtagcacacaccttaacgagagcatagccatagta 1980
 Db 1713 G-----CATAGACAGAGAGCATCCCAATGTC 1739
 QY 1981 taagacacctgtgacctgcaagccatccaaagaatgaagatcgccctaatgtgtglatltt 2040
 Db 1740 TTTGCACCTCTGTACCTGGCAGCATCCAAAGATGAGATGCGCTTAATTTGCTGTATTTT 1799
 QY 2041 attaaatagcgttttaagaatgagagatgtgtactctgagactcaggaacaaatgtcgtg 2100
 Db 1800 AATTAAATTAAGCGTTTAAAGATGAGGTGATCTCGAGATTCAGAGCAATGCTGGGCTT 1859
 QY 2101 aaaggttgtgagagaaagaatgaatcgaatcaggtcgaactgtgtgcatltaatacgaagt 2160
 Db 1860 AAGGTGTAGAGGAGAAAGATGATGATAGAGTGCAGCTTTGTGCAATTTATTACCAAGT 1919
 QY 2161 aaaaaaaggaagtttagctgtatctgttagaatacttaagccaggtgatatgaagcttga 2220
 Db 1920 GAAAAAAGCAAGTTTATGCTGATCTGTAGGACATCTTAGACCAAGGTGATGAAGTCTTGA 1979
 QY 2221 atggaatgagagctatctgaagagagccacatttgaggaagtttacaacatlatctaga 2280
 Db 1980 ATGGAATGGAAGGCTATTGCAAGAGCCACATTTGAGAGAACTTTACAACATTTATCTAGA 2039
 QY 2281 atccaaacctgaaccacaagttgagcttgttltccaagccaaatggagatatctccag 2340
 Db 2040 ATCCAAAGCCTGAAACCAACAGTTGAGCTTGTGTTCAAGCGCATTTGAGATATGCCTAG 2099
 QY 2341 aatacctgtatgacagcatgcacaaactggaaatccagttcagctcagctatgtatccaa 2400
 Db 2100 AATACCTGATAGACACTATGACAACTGGAATCCAGTTCTAGCTCATTTGATATCCAAA 2159
 QY 2401 aatggacgtctccctctatccgttatccctacccaatgagctctgagcagctgagagatgt 2460
 Db 2160 AATGGATCGTCCGCTATATCGGTATCCCTCCATGATGCTCTGSCATGTCAGAGGATGT 2219
 QY 2461 cccgagttctlatcttgagacagcttcoaataaactatggtttgcaaggttgcacaa 2520
 Db 2220 CCCACAGTTCTTATCTGACAGCTTTCAATTAACATAATGTTGCAAGGTTGTCACCA 2279
 QY 2521 gttgagttcaaatlttgagagcagaagatctccctccaaagagatgtagagagccaa 2580
 Db 2280 GTTATATGTTCAATTTTGGAGCAAAAGATCTCCCTTCCAGAGAAAGTGGAGGCGCAAG 2339
 QY 2581 gaatccatgttlaagatltacttccctccagatagaagtgataaataagaagaagaac 2640
 Db 2340 GAATCCTTATGTAAATTTACTTCTTCCAGACAGAAAGTGAATTAATAACAGAGAAAC 2399
 QY 2641 aaaaacagtcagaagaacttggaaacccaatggaaacagagacttcatltaatctcgt 2700
 Db 2400 AAAAACAGTCAAGAAACCTTTGGAACCCAAATGGAACCAAGACTTTTCAATTTATCTCTGT 2459
 QY 2701 ccacgagagagatctcgtgaacgaaatgcttggaatlaaccccttggatcoagctagagt 2760
 Db 2460 TCACCGAAGAGATTTCCGGGAACGAATGCTCGAAATCACCCCTTTGGGAATCAACCTGAGT 2519
 QY 2761 tgcagaagaagagcgaatctcttaggagagatlttaattgaatttgaaacagcttgc 2820
 Db 2520 TCGAGAGGAAGAAAGATTTCTTAGAGAGATTTAAATTTGAATTTGAAACACACTCTGTT 2579
 QY 2821 agatgtagccgcagctgtgtataagctgcagaccatagatgtctccatgtccactccc 2880
 Db 2580 AGATGATGAACCACTGTGTACAACTTCAGACCCATGATGTCTCCTCATTTGCCACATCC 2639
 QY 2881 tggccctcccatatctcccgagagcagctcatalgagagagcccaagcgagagct 2940
 Db 2640 TCACCTTCTCATATATGCCACGAGACAGCTGATGAGAGAGCCACACAGCAAGTT 2699
 QY 2941 gcaaaagtcgaagaagaatagtagcagtgagatgtctgactaagactgcgaagtagagct 3000
 Db 2700 GCAAAAGCTGAAGAGAAATAGTACAGCGAAGTATCGACTATGACTGCGAGGCGCGCT 2759

QY	3001	gggagtagtgcataattatcgcacaatgycgcgcgaatcttcaagctccaagtctgtcgt	3060	
Db	2760	GGGAGTCGTGCATGATTACCGACAGGATGGCGGTGATCTTCAAGCTCCACATTATACGT	2819	
QY	3061	gccaaacaagcatgataataatcattgctccacataaggctccctcatcgagaga	3120	
Db	2820	GCCAAACAAGCATATGTCATCAAAACATTGCTACCATCAAGGCTCTCTCATGAGATAGA	2879	
QY	3121	tgctatagaagaaagatcctatgctgcctagtgccctccctccccaagaaatgtga	3180	
Db	2880	TGTTATAGGAAGAGCATAGGTCAATGCTGGCTAGTGTGCCCTCCCTCCCTCAAGGATGTGGA	2939	
QY	3181	acaagggcacogaaaggacaagtcgtctactatgycatatacaacaatatagccgaatgtga	3240	
Db	2940	ACAGGGGCTTCAGAGGAGACAGTGCTACTGGCCATTACAAATACATTAAGCCGATGACAG	2999	
QY	3241	acaccgtctcatagatgataccactactcttcagataagac	3280	
Db	3000	ACACGGTGCATAGACGAGACCATTAATCTTACAGAGAAGACAGTCATTTCTCACTTACC	3059	
QY	3281	-----	aggagtgtgaagc	3294
Db	3060	TCGCTCTGACACAGGCAGACCAAGTACAGATCAACCAAGGATGGAAGGATTTGTGAAC	3119	
QY	3295	agcagatagacagcatatcacagatccacatcaacaagacaacagcctctccatagaagc	3354	
Db	3120	AGCATATGACAGCATATACAGATCCAGATCAACAAGAACAGCGGCTTCTCTAGAGCG	3179	
QY	3355	gacacacacccgctccaagctctcttgaagctccttgatataaacctcatgtggtctatagcc	3414	
Db	3180	GACCAACACCCCTCCAGATCCTGTGAACGTGTGATACAAACCTCATGAGGTGATGAGCC	3229	
QY	3415	ttcatatagtacgtgaagaatctgcgccctccctcaactgccttatcgaggtcccaacctcg	3474	
Db	3240	TTCAATTAAATGATGTGAAGATCTGCCCCCTCTTACCTGCTTATTCAGAGTCTCACCTTCG	3229	
QY	3475	taccggtctgtccagacaagcccatcaaatctccgcgggaacagcagaaggggcgaca	3534	
Db	3300	CACCTGGCTCTGTCCAGAGACAGTCCGCTCAAGTACTCCGCTMACAGAGCAAGGGCCGCGCA	3359	
QY	3535	gcttccaagagcttccacccaagggaacatctggaagaagaatgtgcatagtatagaagagag	3594	
Db	3360	GCTTCTCAAGCTCCACCAAAAGGGAACACTGTGAAG-----	3395	
QY	3595	aatcgccaatgaacttaacaataaacaacagctagccgagatcagaccacagactga	3654	
Db	3396	-----	-----	3395
QY	3655	gcaagatcacatctcgaaatctgcgtcgaatgtggaatccacatagagggcagataactgt	3714	
Db	3396	-----	-----	3395
QY	3715	ttccaataatccctcggaagtgatgttaagtgtatctgcggtttccaagactagtag	3774	
Db	3396	-----	-----	3395
QY	3775	tgctctcgcttcacgacgacaagctatcgtccgctccaatcagagcgcgagaggaaa	3834	
Db	3396	-----	-----	3395
QY	3835	caggaaatcagtgctcttataatccaaaatgcaaaaacagacagatgggcgtgtcgggaa	3894	
Db	3396	-----	-----	3395
QY	3895	gaacttgaccaaagacacacagatcagctggagacatgtgtctcatctggagaagaatgacgg	3954	
Db	3396	-----	-----	3395
QY	3955	cagcgagctccgacactgcagctggcgcccttggttaccagctgacgaagaagcgcgactag	4014	
Db	3396	-----	-----	3395
QY	4015	catlgyggccaatgylagctatgttgtctctcaagaaaagtgcagtgccctca	4074	

	Db	3396	-----	3393
QY	4075	actcagcccaaacccgagagagaggtlaaagaactacgagacgtgttccagaagaacacgga		4134
Db	3396	-----	AGGAGGAGGTAAAAAGCTACGGAGCAGCTGTCCAGAGGAAGCACGGA	3440
QY	4135	gaccgggtctagcaagtgtgagatgtgaagaaactgatacgcgcgcagagccagcggaattccac		4194
Db	3441	GACCGGGCTAGCAGTGTGAGATGAGGAACCTGGATGACAAAGACAGGCCGCGGGGAAATCCAC		3500
QY	4195	aaatgcacacatgaacaagctatagctcgcgaaggaatactgacttcccttggtggccgct		4254
Db	3501	GGATGGCAGCATATGAAAGATTACAGCTCAGAGAAATCTATCTTTCTTG6GGTCCGCT		3560
QY	4255	ggcctctgacagccagcttcaagtatttccctggaatggcccttgccctctcagctagtgag		4314
Db	3561	GGGCTCTGACAGCCAGCTTACGAGCATTTTCCATTGATGGTCTGGGCGCTCTCAGCTAGTGGG		3620
QY	4315	acgcagagacccttggtactctcctgcaatgggtgacatcagttgggaatgatatgaataaa		4374
Db	3621	ACGCCAGACCCCTGGCCACCTCTCAATGGGTGCAATTCAGGTGGGGATGTGTGTAAGAA		3660
QY	4375	gggacacgtctggaagttagaaatcaatccggggcgccgacctgtgtgtaaacacagtttcaa		4434
Db	3681	GGGACACGCTGGAGGTAGAGATCATCCGGGCTCCAGGCGCTTGATTAACACAGTTCMA		3740
QY	4435	gaacatgccaagcacglatgtcaaggtgtlatctgttgaacaacggaatctgtcatagccaa		4494
Db	3741	GACACTGCCAGCACCTATATGTCAAGGTGTACCTGTGTAGACAAGGAGTCTGCAATAGCCAA		3800
QY	4495	aaagaataacgaagtgtgcgagaagaagacctggagccctgttaccagaagctctgttctt		4554
Db	3801	AAAGAAACCAAGGTGGCCGGAAGAAGCGTGGAGCCCTGTATCAGGACGTGTATCTTT		3860
QY	4555	cgaagagaccccccaaggggaggtgttacaagatcatgtctggggagaattatggtcgat		4614
Db	3861	CGAAGAGAGACCCCCAGGGAGGAGGTATTACAGATCATTTGTGTGGGGAATTTATGGCCGAT		3920
QY	4615	ggatcaacaatctcttattgaggagtgagccagatactctatagatgaactggaacttccaa		4674
Db	3921	GGATCCAAATCTTTATGGAGCTGGCCCGCATACTCTTGATGATGAACCTGGAAACTATCA		3980
QY	4675	catgtgatttggatgtgttcaaaactcttccctcttccctcctcctagtataccaactcggc		4734
Db	3981	CATGGTAATTTGGATGGGTCAAACTCTTCCCTCTCTCCCTCATGTATATCAACCTTTGGC		4040
QY	4735	acctctgacaagaagagcttccccaatcgtctctcggaaagttcttcccgagaccttcttacc		4794
Db	4041	ACCTCTGACAAAGAAGAGCTTCCCAATCGTCTGTGGAAAGTTCTACTGTGACCTTTACTC		4100
QY	4795	tgttctatagcaactataaactgttgttccaacaacacagcagatatacaaaaaccagaaga		4854
Db	4101	TGTTCTATGCAACTATATAAACCTGTTGTCTCAACAAACACACGTTTACAAAAACCGAAGAA		4160
QY	4855	aaacgcacagtgtagaagccctctgttaacacagcagatctcttgaatgtgttcttacaagaagcca		4914
Db	4161	ACCACCCAGGTGG- GGAACCTGTGTCACACTGCAATGCTTGATGTGTGTCTTCGGAGCCCA		4219
QY	4915	cgttctagggatatacaacagcagcttccgttgttctcagaaggaag----tcgtacaacattgtgc		4970
Db	4220	CGTCTAGGGATTAAGAAGCATTCTCTGTCCTCTCAGAGGAAGCAATCCGTGCACAGTGTGCC		4279
QY	4971	ctagcaagaag 4980		
Db	4280	CTAGCAAAAG 4289		
RESULT	8			
LOCUS	AF199326	5031 bp	mrna	04-JUL-2000
DEFINITION	Rattus norvegicus RIM2-3b (Rim2-3b) mRNA, complete cds.			
ACCESSION	AF199326			

RESULT	8				
AF199326					
LOCUS	AF199326	5031 bp	mRNA	ROD	04-JUL-2000
DEFINITION	Rattus norvegicus RIM2-3B (Rim2-3B) mRNA, complete cds.				
ACCESSION	AF199326				

```

VERSION      AF199326.1  GI:8925865
KEYWORDS
SOURCE       Norway rat.
ORGANISM     Rattus norvegicus
              Chordata: Craniata: Vertebrata: Euteleostomi:
              Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:
              Rattus.
REFERENCE    1 (bases 1 to 5031)
AUTHORS      Wang, Y., Sugita, S. and Sudhof, T. C.
TITLE        The RIM/NIM family of neuronal C2 domain proteins. INTERACTIONS
              WITH RAB3 AND A NEW CLASS OF SRC HOMOLOGY 3 DOMAIN PROTEINS
JOURNAL      J. Biol. Chem. 275 (26), 20033-20044 (2000)
MEDLINE      20347919
REFERENCE    2 (bases 1 to 5031)
AUTHORS      Wang, Y. and Sudhof, T. C.
TITLE        Direct Submission
JOURNAL      Submitted (27-OCT-1999) Center for Basic Neuroscience, The
              University of Texas Southwestern Medical Center, 6000 Harry Hines
              Blvd., Dallas, TX 75235-9111, USA
FEATURES
source       1. 5031
              /organism="Rattus norvegicus"
              /db_xref="taxon:10116"
gene         1. 5031
              /gene="Rim2-3b"
              /gene="Rim2-3b"
              /codon_start=1
              /product="Rim2-3b"
              /protein_id="AA01648.1"
              /db_xref="GI:8925866"
              /translation="MSAVLGRGRPAPPPASQPPPEPMDLSLTFEERKTIQAVM
              /DROKKEEKEDSVLKLHQDFEMKKEYKKKEESQDQEQDGPCTCTCKTKRAD
              CGCHNCYCCOTKFCARGCGRVLSNKNVMMVNCILCKQOEILTKSGAMVNSGNTPO
              OPDOKALRGLESEAPQEKAKLHEOTQFCGPEDGSYPAYVRGAHGLTRDSDIKNG
              SGKMHQIASDMPDRKRSYSRDNRREYDSEEESEYSOYVSESTMRSESDYADR
              RSQREPFYEPEDHLNYSRDSNRGRHRSKEYIVDEEDVESREYESTMRSESDYADR
              SDEMLARTPVKQPYEEQMRTIAEYRSARHERHSRHSVLSANLEDSRSLRIMDRPS
              RQVRSVSRRAAMENQRTSMETRAQCSSTPQRTNHSPTPPRSPIDRDELR
              ADSLRKHHLDPSSSAVRTKREKMETMLRNDLSQDSQSVPPPRPRKSKGGKMR
              QVLSSESEELAPTEYSCDVELESESESGKGMEXSMLHASMSSEASPSLAP
              VTWQSKGDLRIGRLILNKLKDGSPVDSGAMGLKVVGMESGSLCAFTLVK
              KGSADPYGHRPGDEVLNMGRLLOGATFEELYNIILSKPEQVELLYVSRIGDP
              RIPTDPAQLSSSSSFESOKMDRPSIVTSPMSGMGLADVQSLGSLIKLMPKY
              GHQILVITLGAKLDPREDGRPNRYKTYFIPDSDKRKRTKYKTKLEPWNQTF
              IYSVHRREFRMLLEITLMDQAVRSEESFLEGLLELLETALDEDEHWYKQTHD
              VSSLPHPSPYMPRLQHGSPDRRLQSKRISDSVSDVDCEDGVAVSDYRHQK
              DLOSSTLSVPOVSNHSPSGSPHRYVIGRTSRWSPSVPPQNNVDQGLGTBAT
              GHYNTISMDRHYVDHYSSESDSHFTLIPSRHROTSEHHRRDGRDEADARQYH
              RSRSTEQPRLIERTTSTRSRADTNLMRSMPSLMTGRSAPSPALSHPMTGSVO
              TSPSTPYTGRRGRLPOLPRKGTLERGGKRLRSVORSTETGLAVEMRNMTROAS
              RESIDGSNYSVSSSEGNLIPGVRLASDSQSPFLDGLGROLVGRQTLATPAQDIOY
              GMDOKGQLEVEILIRAGLVYKPGSKTTPAPYKVTLLNNGVIAKKTKIVAKTLEP
              LYQOLSFSEESPOGVLIQIIVWDYGRMDHKSEMGVAQILDLLELSNMVIGFKLEP
              PSSLVDPITLPTRRASQSSLESSTGSPYSRS"
BASE COUNT  1503 a 1166 c 1246 g 1116 t
ORIGIN
Query Match      57.4%; Score 2860.4; DB 94; Length 5031;
Best Local Similarity 78.0%; Pred. No. 0;
Matches 3938; Conservative 0; Mismatches 261; Indels 851; Gaps 7;

```

```

Db 141 CCACCTCAGGAGAGAGAGAGAAATCATCCAGGCTGTCTAGTGCATGAGAGAAAGA 200
Oy 181 agaggagaaggagcagtcgctgtcctaaagatccaaagaacacaaagcacaccgacaca 240
Db 201 AGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227
Oy 241 gtagtcccttagtgaggatcaactgtaataacgltcgtgagcccgacaa 300
Db 228 ----- 227
Oy 301 acaaccaatgagaagagaccaccaagaagctgacaccacaattgataataga 360
Db 228 -----AAAGCTCATCAACAATTTGAAATGTATTAAGGA 260
Oy 361 gcaagtcaagaagtgtgagaagagatcgagcagcagcagaagcagaagggtgagcccc 420
Db 261 GCAAGTCAAGAAATGGAGAGAGATCAACACAGCAGCAAGCAAGAGAGAGAGAGAGAGAG 320
Oy 421 gacctgagcatctgcacacaagaacaatttcagatgagatgagcgacaaatgttcccta 480
Db 321 GACCTGGCGCATCGCCACAAGCAAAATTTGCAGATGATGCGGCGCAWAACTGTTCAATA 380
Oy 481 ttgcacaaccaagttcgtgctcgtatgtgagagtgctgagtgctttagcttaaaacaagt 540
Db 381 TTGCCAAACCAAGTTCTGTGCTCTGTGGAGGTGAGTGTATTAACGCTCAACAAGGT 440
Oy 541 tatgtggtgtgataattgtgtccgaaacaacaagaatccctactaaatcagagagcatg 600
Db 441 TATGTGGTGTGATATTGTGTCCGAAACCAACAGAAATTCCTACTAAGTGGCGCGGTG 500
Oy 601 gtttataatagtggtctaaacacatcgcagcaacctgatacaaaagttcctcgaggct 660
Db 501 GTTTTAAATATGTGGGTCTAATACACCCAGCGCTGTATCAAAAGGCTCTTCGAGGGCT 560
Oy 661 tcgaaatgagaagccctcctcaggaagaagaacaaacaaacagcagagcccgatcca 720
Db 561 TCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 620
Oy 721 aggaagcccgatgactatcagatcagtcagtcagtcagtcagtcagtcagtcagtcagtc 780
Db 621 GCGAGCCCGGTGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680
Oy 781 aagacagatcatttaaataatgatacgaagatggaagcccgatgagtcagtcagtcagtc 840
Db 681 AAGACAGATTTCTAATTAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 740
Oy 841 ttgcagcagaagaagaagatccatcagtggtcagaggaatcaaaatcgaagatacgaagaag 900
Db 741 TTCACAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 800
Oy 901 tgaagaagaagaagatcactcagatgagttccttcagatggtlacaatgccaaagatctcc 960
Db 801 TGAAGAGAGAGAGAGAGATTTACAGATATGTCCTTACAGATAGACAAATGCTTGAATCTCC 860
Oy 961 ttgcagatgctgataagcagatctcagtgtagcctcaatttatgaagaacctgtgca 1020
Db 861 ATCAGATTAATGCTGATAGAGAGATCTCAACGTGACCTCAATTTTATGAGAGAACCTGATCA 920
Oy 1021 ttaataatcagagatctcaacagagagagccatagacattccaaagatataatgttga 1080
Db 921 TTTAATTTTATGAGGATTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 980
Oy 1081 tgaatgaatgtgagagagagatgaatatgaagaagaagaagaagaagaagaagaatgca 1140
Db 981 CGACGAGAGATGTGAGAGAGAGAGAGATGAATGAAGAACAAGAGAGAGAGAGAGAGAGTACCA 1040
Oy 1141 ggcagcgtcaagaagtgaatccaaatctggcccggtatcccggttaaagccaaacctcaga 1200
Db 1041 GGCAGCGCTACAGAGATGATCAAAATTTGGCCGCGTATCCGCTAAAGCCACAACCTTATGA 1100
Oy 1201 agaacaaatgctgcatccagctgaggtgtccagagcagacatgagagaagaagcagatga 1260
Db 1101 AGAGCAAAATGCGGATTCACAGCTGAAGTGTCCGCGGACGACAGACAGAGAGAGAGCAGAGTGA 1160

```

QY	1261	tgtttccttggcaaaacgcctgaaactagaagaattccaaagatttccctctgtaagagtgtag	1320
Db	1161	tggttcttttgcaaaatgcctgaacttgaamaattccaaagatttctctgtaagatgtag	1220
QY	1321	accatcaagagcaaaagatctgltacgtgaagtagagctgcgaatgaaaccaaagatctgta	1380
Db	1221	accatcaagagcaaaagatctgtrtctgaactgtagagctgcgaatgaaaccaaagatctgta	1280
QY	1381	ttcaatggaagaagactcgaagagctcgaaggacaagaattcttalcacaagaagactcaaa	1440
Db	1281	ttccatggaagaagaactcgaagagctcgaaggacaagaattcttalcacaagaagactcaaa	1340
QY	1441	tcatagctctccaccacccctcgagcgagacccttaccgcttgtagaccagaactgaagcg	1500
Db	1341	tcataagctctcttacccttaccgctcgagcgagacccttaccgcttgtagaccagaactgaagcg	1400
QY	1501	cgctgactccctacgagaaacagcaccaacttagatcccagctctgcgtctgtgagaacagaa	1560
Db	1401	tgccacatccccctacgagaaacagcaccaacttagatcccagctctgcgtctgtgagaacagaa	1460
QY	1561	gcgagaaaaaatggaacacattgttaagaaatgatcttttgaagtcagacaagctcgagtc	1620
Db	1461	gcgagaaaaaatggaacacattgttgagaaatgatcttttgaagtcagacaagctcgagtc	1520
QY	1621	agtgaagccgcgcccccaacgaagccctcatalaatccaagaagaagaggtlaaaatctgcacagt	1680
Db	1521	agtgagccgcgcccccaacgaagccctcatalaatccaagaagaagaggtlaaaatctgcacagt	1580
QY	1681	ttcaactgacagctctcgagaggaagcttgagcatccacacactgaatlatacaagctgtgtga	1740
Db	1581	ttccactgacagctctcagagaggaagcttgagcatccacacactgaatlatacaagctgtgtga	1640
QY	1741	tgtagagctgtgaaagcgsagagtggtgtagttagaagaaggcgagtcacaaaggaaaaagaa	1800
Db	1641	cgtgagagttgaaacgagacgagcgttagtgtagaagaaggcgagtcacaaaggaaaaagaa	1681
QY	1801	aactagtgagcagaggagatttltgctgattctaacacaaagctctgagagacaaaagaag	1860
Db	1682	-----	1681
QY	1861	gatgtactactgltgagcaactctttggaagaagatttgaatggtctcgagctcagaattaa	1920
Db	1682	-----TGAGTTCACAGCTGTTGGAGCATGCGTCTTG	1712
QY	1921	ggaactctgggttagatctactgtgtagtgaacaacaccttaacgagagatcagtagtga	1980
Db	1713	G-----CATGACAGTGAAGCATCCCCAATGTC	1739
QY	1981	taagcaaccctgtgacccctgagcagccalcacaagaatgtagatctgccttaattgtcgtatct	2040
Db	1740	ttttgacacccctgtgacccctgagcagccalcacaagaatgtagatctgccttaattgtcgtatct	1799
QY	2041	atthaataagcgtttaaaagaatgggagtgtagctcgagactcaagagcgaaatgctcgggct	2100
Db	1800	attthaataagcgtttaaaagaatgggagtgtagctcgagactcaagagcgaaatgctcgggct	1859
QY	2101	aaaggtcttagaggaagaagatctcgtgaacccaggtcgcaacttctgcatttataccaaga	2160
Db	1860	aaaggtcttagaggaagaagatctcgtgaacccaggtcgcaacttctgcatttataccaaga	1919
QY	2161	aaaaaagaagaattagctgtagatctgttagacatccttagacaagctgtagaagctctgga	2220
Db	1920	gaaaaaaagaagatttagctgtagatctgttagacatccttagacaagctgtagaagctctgga	1979
QY	2221	atggaatgaggagagcattgtcgaagagggccacatttgaggagagtttacaacattctttaga	2280
Db	1980	atggaatggaagagcattgttgcgaagagggccacatttgaggagagtttacaacatttattcttaga	2039
QY	2281	atccaacactgaacacaaagttagagctgtgttlttaagaagccaattggaataattcctag	2340
Db	2040	atccaagagcctgaacacaaagttagagcttctgttttcaagagccattggaataatmgcctag	2099

[illegible]

```

|||||
Db 3180 GACCCACCCCTCCAGATCCTCTGAAGCTGCTATACAACTCATGAGTCGATGCC 3329
|||
Qy 3415 ttcaatlaactggaagaactgccccctccctcaactgacctatcgagctcaacctcg 3474
|||
Db 3240 TTCAATTAATGACTGGAAGATCTGCCCTCTTCACTGCTTATCGAGTCACTCCGCG 3299
|||
Qy 3475 taacgggtctgtcccgacaaagcccaactcaagtctccgggaaacagagcgagggcgaca 3534
|||
Db 3300 CACTGGCTCTGTCCAGACAGTCCTGTAAGTACTCCGTTACAGAGAGAGAGGGGCGGCA 3339
|||
Qy 3535 gcttcacagcttcccaaaaggaacattggagagagagtgctatagatagagagag 3594
|||
Db 3360 GCTTCTCAAGCTCCACCAAGGAACACTGMAAG----- 3395
|||
Qy 3595 aaatgcacaatlgaaaactaacaatatacaacaggtagccggatcagaccccgactgga 3654
|||
Db 3396 ----- 3395
|||
Qy 3655 gcaagattacccattcgaaatgcgtccgagatggtgatacacaatagagggcagatactgt 3714
|||
Db 3396 ----- 3395
|||
Qy 3715 ttccactaaatccctcgagacagtgatgtaagtgtatctgcgtttcaggagactagtag 3774
|||
Db 3396 ----- 3395
|||
Qy 3775 tgcctcgtttcagcagcacaagctacatgctccgccaatcagagcgccgagagagaa 3834
|||
Db 3396 ----- 3395
|||
Qy 3835 caggaaaatcagtgcttcttaacatcaaaatgcaaaacagacagatggcgctgtcgggaa 3894
|||
Db 3396 ----- 3395
|||
Qy 3895 gaacttgaccaaaagcaccagcatcagtgagagacatgtgctcaactgagagaagaaatgacgg 3954
|||
Db 3396 ----- 3395
|||
Qy 3955 cagcagctcgcagactgcagtgggcgccctgtggtaacagtgcgaagaagcgcgactag 4014
|||
Db 3396 ----- 3395
|||
Qy 3995 cgttcagagataccaagcgactgtgtctcctcagcgaagaaagtgcagctctca 4074
|||
Db 3396 ----- 3395
|||
Qy 4015 cattggggccaataatgtagctatgtgtgtctcctcagcgaagaaagtgcagctctca 4074
|||
Db 3396 ----- 3395
|||
Qy 4075 actcagccaaaccgaaggaggtaaagaactacagagacatgttcaagagaacacagga 4134
|||
Db 3396 -----AGGAGAGGTAAAGCTACGACGACGCTCTCCAGAGAACGACGGA 3440
|||
Qy 4135 gaccgggtagcagtgagagatgagaaactgatatgaccccgccagccagccgggaatccac 4194
|||
Db 3441 GACCGGGTAGAGCTGAGATGAGAACTGGATGACAAAGACAGCGCCGCGGATCCAC 3500
|||
Qy 4195 aagatggcaacatgaacagctatagcttcogaaagaaatctgattctccctggggtccgct 4254
|||
Db 3501 GGATGGCAGCATGAACAGTTACAGCTCAGAGGAATCTGATCTTCTGGGGTCCGCT 3560
|||
Qy 4255 ggcctctgacagcagcttcaagtattccctggatggcctgggcccctgctcagcagtgagg 4314
|||
Db 3561 GGGCTTGACAGCGCAAGTTCAGCATTTCCCTTGAGTGCTGGGGCCCTGCTCAGCTAGTGGG 3620
|||
Qy 4315 acgacagacccttgactctctgcaatgggtgacatcgaatgggtgggaatgatatgataaaa 4374
|||
Db 3621 AGCGCCAGACCCCTGGCGACTCTCTCAATGGGTGACATTCAGGTGGGATGATGAAGAA 3680
|||
Qy 4375 gggacagctgggaagtgaataatcaccggcgccgagcctgtgtgtaaaacaggtttccaa 4434
|||
Db 3681 GGGACAGCTGGAGGTAGATCATCCGGCTCGAGGCCCTTGATGAATAAACAGGTTCCAA 3740
|||
Qy 4435 ggaacagcagacccgtatctcaagtgtagtctgttagacaagagagctcatagccaa 4494
|||

```

```

Db 3741 GACACTGCCAGCACCATATATGTCAAGGTGTACTGTTAGACAAACGAGATGTCATAGCCAA 3800
|||
Qy 4495 aaagaaaacccaggtggcgagagaagacccttgagagccctgtaccagacgtctgttcctt 4554
|||
Db 3801 AAAGAAACCAAGGTGGCGAGAAAGAGCTGGAGGCCCTGTATACAGCAGCTGTATTCTT 3860
|||
Qy 4555 cgaagagagcccccgaggaggtgttaccagatcatgtctgtcggagatattgtcgtat 4614
|||
Db 3861 CCAAGAGAGCCCCGAGGGGAGAGGTATTACAGATCATTTGCTGGGAGATATATAGCCGAT 3920
|||
Qy 4615 ggaacacaatcccttlatgaggtggccagabactcttagatgaactggaactatccaa 4674
|||
Db 3921 GGATCAAAATCCTTATATGGAGTGCGCCAGATACCTTATAGATGAAGTGAAGTGAAGTGAAG 3980
|||
Qy 4675 catggtgatgtagtggttcaaacctctccctccctccctccctccctccctccctccctcc 4734
|||
Db 3981 CATGGTAAATGGATGGTTCAAACTCTTCCCTCTTCCCTCTTCCCTCTTCCCTCTTCCCTCTT 4040
|||
Qy 4735 acctctgacaagaagagcttcccaatcgctctctggaaggttccacggacctcttactc 4794
|||
Db 4041 ACCTTGACAAAGAGAGCTTCCCAATCGTCTGGAAGTTCAGTACGACTTCTTACTC 4100
|||
Qy 4795 tgcctaatagcaactataaaactgtgtgtcacaacaacccagcgagatatacaaaacagaagaa 4854
|||
Db 4101 TCGTTCATAGCAACTATTAAGCTGTGTGTCACAAACACACGCTTACAAAAACAGAGAA 4160
|||
Qy 4855 aagcacaagtggaagccctgtgtaacactgcatgtgttgatgtgtgtgtgtgtgtgtgtgt 4914
|||
Db 4161 ACCACCAGAGTGG-GGACCTGCTCCACGCAATGCTGATGTTGTGTCTCCGGAGCCCA 4219
|||
Qy 4915 cgtctagagataccaagcgactgtgtgtctcctcagagaa-----tcgtacaactgtgccc 4970
|||
Db 4220 CCTCTTAGGATTAAGAGCCATCCTGTGCTCTCAGAGSAGAACCATCCGTGCACAGTGTGCC 4279
|||
Qy 4971 ctagcaaaagg 4980
|||
Db 4280 CTAGCAAAAG 4289
|||

RESULT 9
AF199327 LOCUS 5031 bp mRNA ROD 04-JUL-2000
DEFINITION Rattus norvegicus RIM2-4A (Rim2-4A) mRNA, complete cds.
ACCESSION AF199327
VERSION AF199327.1 GI:8925867
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 5031)
AUTHORS Wang, Y., Sugita, S., and Sudhof, T. C.
TITLE The RIM/NIM family of neuronal C2 domain proteins. INTERACTIONS
WITH rab3 AND A NEW CLASS OF src HOMOLOGY 3 DOMAIN PROTEINS
J. Biol. Chem. 275 (26), 20033-20044 (2000)
JOURNAL
MEDLINE 20347919
REFERENCE 2 (bases 1 to 5031)
AUTHORS Wang, Y. and Sudhof, T. C.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The
University of Texas Southwestern Medical Center, 6000 Harry Hines
Blvd., Dallas, TX 75335-9111, USA
FEATURES
source
1..5031
location/Qualifiers
1..5031
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..5031
/gene="Rim2-4A"
52..4110
/gene="Rim2-4A"
/codon_start=1
/product="RIM2-4A"

```


1681 ttcaactagcagctcgagagagagctggtcattccacacctgagatatacaagctgtgtatga 1740
1581 TTCACGTGAGCAGCTCAGAGGAGAGAGTGGCATCCAGCCTGAGTATACAGCTGTGATGA 1640
1741 tctgagactgtaagaacggaagatgtgtgagtgtaagaagggagacatgcaaaaaggaaagaaa 1800
1641 CGTGGAGATTGAAAGCCGAGACGTAGTGAGAAAGGAGACA----- 1681
1801 aactagtagagagagagattgttcgattcttaacacaggtctgtagagagaaaagaag 1860
1682 ----- 1681
1861 gatgtactatgtgtgccaactcttggaaaggatttggaaatggtctgagccctcaagattaa 1920
1682 -----TGAGATGACAGCTGGTTGGAGACATCGCTTGG 1712
1921 ggaactcggggtagatatactctgtatgagcaaacaccttaacgaggaacatagacatgtga 1980
1713 G-----CATGACAGTGAAGGCATCCCCAATGTC 1739
1981 taagcaacctgtgacctggtgcaagccattccaaagatgtgagatcgccctaattgttcgtattt 2040
1740 TTTGCACCCCTGTGACTGGCACCATTCCAAAGATGAGATPGCCTAATTGGTGTATTTT 1799
2041 attaaataagcgtttaaaagatggagatgttacctcgagactcaagagcaatgctggtgt 2100
1800 ATTAAATAGCGCTTTTAAAGATGGAGTGTACCTCGATTCAGAGCAATGCTGGGCTT 1859
2101 aaaggctttagaggaagaagatgactgaatcaggtcagacttgtgcatttataccaaagt 2160
1860 AAAGGTTGTAGAGAGAAAGATGACTGCAATCAGTGCACCTTTGTGCATTTTACCAAGT 1919
2161 aaaaaaggaagtttagctgtatctgtaactgttagagacatcttaagccaggtgataagcttga 2220
1920 GAAAAAGGAAGTTTAGCTGATGACTAGTACATCTAGACATCTTAGACCAAGTGATGAAGCTTGGGA 1979
2221 atggaatggagagcatttgcgaaggagccacatttggaaagatttaacaaattctcaga 2280
1980 ATGGATGGAAAGGCTATTGCAAGAGAGCCACATTGAGAGAAATTACAACTTATTTCTAGA 2039
2281 atccaaacctgtaaccaaaagttgagcttgttctcaaggccaatttggagatatctctag 2340
2040 ATCCAAAGCCTGAAACCAACAGTGTGAGCTTGTTCAGAGCCGATTTGAGATATAGCCTAG 2099
2341 aatacctgtatgaacgcatgacacaaactgtgaatccagttctagctcatcttgaatctcaaaa 2400
2100 AATACCTGTATGACACATGACACAACTGGAATCCAGTTCTAGCTCATTTGAATCTCAAAA 2159
2401 aatgagcgtctctctctatatactgcttaccctcaacaaatgagtcctgtgacatgtgtg 2460
2160 AATGGATGCTCCGCTTATATTCGTTTACCTTCCCATGATAGTCTTGCGCATGCTGAGGAGATGT 2219
2461 cccgagcttctatcttgcagacagcttccaataaaactatggttgcacaaggttgcacaa 2520
2220 CCCACAGTCTTATCTGAGACAGCTTCAATAAATACTATGTTTGTGACAAAGTTGGCTACCA 2279
2521 gttgatagttacaatttttggagcaaaagatctcccttcaagggaagaatggagagcaag 2580
2280 GTTAAATAGTTTACAAATTTTGGAGCAAAAGATCTCCCTTCCAGGGAGATGGAGGCCAAG 2339
2581 gataccttctgttaagattacttctctccacagatagaagtataaaaaataagagaagaac 2640
2340 GAATCTTATGTTAAATTTTACTTCTTCCAGACAGAAATGATATAAACAAGAGAGAAC 2399
2400 AAAAACAAGTCAAGAAAACTTTGGAACCAATGGAACACAGACTTTCATTTATCTCCCTGT 2459
2701 ccaaccgaagaagaattcgtgaacgaatgctggaaatataccttggagatacgaagctagagt 2760
2460 TCACCGAAGAAGATTCGGGAGCAATGCTGAAATGCACTCTTGGGATTCAAAGCTCGAGT 2519
2761 tcgagaagaagagagcaattctttagagagatttaattgaattggaacagcttctgt 2820

2520 TCGAGAGAGAAAGATGATATTCCTTAGAGAGATTTTAAATGAAATTTGAAACAGCTCTGT 2579
2821 agatgatagagccgacatggttataagctgtgaagaccatgatgtctcccaattgccaactccc 2880
2580 AGATGATGAACCCACACTGTGTACAAACTTCAGACCATGATGTCTCTCATATGCCACTCCC 2639
2881 tgcgccttcccataatctgccccgagagcagctccaatgtgagagagcccaagcgcagagct 2940
2640 TCACCTTCTCCATATATATGACCAAGAGACAGTGCATGAGAGAGCCCAACACGCAAGTT 2699
2941 gcaaggtcgaagaagaataagtgacagtgaaagtgtctgaactcagactcgagagatggcgt 3000
2700 GCAAAAGTTCGAAGAAATAGATGACAGCAAGTAATCCAGCATGTAGCTGGAGAGACGGCT 2759
3001 gggagtagtgcagatatacgacacaaatggccgagatcttcaaaagctccaagttgtcggt 3060
2760 GGGAGTGTGTACAGTATACCGCACAGATGGCCGTGATCTTCAAAAGCTCCACATATACAGT 2819
3061 gccagaagaagtcataatgtcataaatcattgtccacatcagatagggtctcccatcgataga 3120
2820 GCCAGAAACAGTCATGATCATCAAAACCATTTGCTCACCATATAGGCTCTCTCATCGAGTAGA 2879
3121 tgtataaggaagacaagatcgtgtgcgttagtgccttccctctcctcaaaaggaatgttga 3180
2880 TGTATATGAAAGAGATAGTATAGTATGCTGCGCTAGTGCCCTCCCTCCCAAGGAATGTGA 2939
3181 acaagggacacgagagacagctgtctactgtgccaattacaacaaatlaagccgatgtatg 3240
2940 ACAGGGGCTTGGAGGAGACACGCTACTGCGCATTAACAATTAACATTAAGCCGATVGSACAG 2999
3241 aacacgtgtcaatgataagtcacactactctctcagatagagac----- 3280
3000 ACACGGTGTATGACACGACCATTTACTCTTCAGAGAGACACAGCATTTTCTCATCTCTAC 3059
3281 -----aggattgtgaagc 3294
3060 TCGCTCTGACACAGACAGACCATGAGCATCACACAGAGATGAAAGGATTTGTCAAGC 3119
3295 agcaaatagacagccatattcagaagatccagatccaagaagaacagccctctctagaagc 3354
3120 AGCAATGATGACGCCATATACAGATTCAGATCAACAGAACAAAGGCTCTCTTAGAGCG 3179
3355 gaccacacacgcgtccagatctcctctgaacgtcctgaatacaaacctatgagatgcatgccc 3414
3180 GACCAACACCGGCTCCAGATTCCTGTAACGTGCTGATPACAAACCTCATGAGTGCATGCC 3239
3415 ttcaatlaatgacttgaagaatctgcccctcccaactgacttctgaagtctcaacctgt 3474
3240 TTCATTAATGACTGGAAGATTCGCCCTCTCCCTTCACCTGTGATGAGGTCTACACCTCG 3299
3475 taacgggtctgtccagaagaagcccatcaagttactccggagacagagagagggccgaca 3534
3300 CACTGGCTCTGTCCACACCAAGTCCGTCAAGTACTCCCGTAAACAGAGAGAGGGGCGGCA 3359
3535 gcttccacagcttcccaacaaaggaacatgtgagagaagtgctatgatatagaagagag 3594
3360 GCTTCTCAGCTCCACCAAGGAGACACGTGCAAG----- 3395
3595 aaatgcacaatgaaacttaacaataatacaaacagtgtagccgatatgaacccagactgga 3654
3396 ----- 3395
3655 gcaagatlaacattggaaglatcgtccagatgagatccacatagagggcagatactgt 3714
3396 ----- 3395
3715 ttccactaaatcctcggacagtgatglaagtgatgatatctcggttccaagagactag 3774
3396 ----- 3395
3775 tgccttcgttttagcagacaaagctacatgctcgttccaatcagagcggccgagagga 3834

Db 3396 ----- 3395
QY 3835 caggaaatcagtgctttatccaatgcaaaacagacagatggtcggtcg99aa 3894
Db 3396 ----- 3395
QY 3895 gaacttgacaaaagcaccagcatcagtgagacatgctcacttgagagaatgacg 3954
Db 3396 ----- 3395
QY 3955 cagccagctccagacacgcagtggtcgccctggtaccagtggtgcaagaagcggtacag 4014
Db 3396 ----- 3395
QY 4015 cattgggccaatgtgtagctatgtgtgctctcaccgaaaagtcgcagtgctctca 4074
Db 3396 ----- 3395
QY 4075 actcagccaaccgaaggaggtaaaaaagctacgagcactgttcagagaagcagga 4134
Db 3396 ----- 3395
QY 4135 gaccgggtcagcagtgtagatgtagaactgtgacccgcaggtccagccggaatccac 4194
Db 3441 GACCGGGTGAAGAGTGAAGTGAAGAACTGATGACAAAGACAGCCAGCCGGGAATCCAC 3500
QY 4195 agatggcagcatgaacagcagctatagctcgaaagaaatctgactctccctgggtccgct 4254
Db 3501 GGTGTGCACATGAACAGTTACAGTTCAGAAAGAAATCTGATCTTCTTGCGGCTCCGCT 3560
QY 4255 ggcctctgacagcagtlcagtgatctctctgtagtggtcgccctgctcagcagtlggtg 4314
Db 3561 GGGCTGTGACAGCAGCATTCAGCATTTCCTTGATGTCGTGGCCCTGCTGACGTAGTGGG 3620
QY 4315 aagcagacccctggtcactccctgcaatgggtgacatccaggtggaatgtagataaa 4374
Db 3621 AGCCGACGACCTTGCGCATCTCTCAATGGGTGACATTCAGGTGGGATGATGAAGA 3680
QY 4375 gggagcagctggaagttagaaatcaccggtcggtcggtcgtgtgttaaaccaggttcca 4434
Db 3681 GGGACAGCTGGAGGTGAGATTCATCCGGGCTCGAGGCTTGTAGTAAACCAAGTTCCAA 3740
QY 4435 gacactgccaagcagctatgtccaagtgtatctgttagaacaacgagtlcagcatagcca 4494
Db 3741 GACACTGCCAGCAGCATATGTCAAGGTGTACGTGTAGACAGGAGTCTGCATAGCCAA 3800
QY 4495 aagaacaccaaagtggtgcgaagaagccctgtagccctgtaccagcagctctgtcctt 4554
Db 3801 AAGAAACCAAGGTGGCGAAGAAAGAGCGTGGAGCCCTGTATCAGCAGCTTATCTTT 3860
QY 4555 cgaaggagagccccaagggtgtttacagatcatgtctggtggagatagtgctgat 4614
Db 3861 CGAAGAGAGCCCGGAGGGGAGGTGTACAGATCATGTCTGGGGAATVATGCGCGAT 3920
QY 4615 ggaatcaaaatcctttatggagtggtgcccagatactctatagatgaactggaactacca 4674
Db 3921 GGATCACAATCTTTATGGAGTGGGCCAGATACCTTATGATGAACGGAACATATCCAA 3980
QY 4675 catgtgattgtgattcaaacctctccctcctcctccctagtagatccaactcggc 4734
Db 3981 CATGGTAATVGTGGTCAAACTCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4040
QY 4735 acctctgaaagaagaggtctcccaatcgctctctgaaaagtcttaccggaactcttactc 4794
Db 4041 ACCTCTGACAAAGAGGCTTCCCAATCGTCTCTGGAAGTTCTACTGGACCTTCTTACTC 4100
QY 4795 tctgttcatagcaactataaaactgtgtgccaacaacacagcagatatacaaaaacagaaga 4854
Db 4101 TCGTTCATGACCACTATATAAACTGTGTGACACAAACACAGGTTACAAAAACAGAAAGAA 4160
QY 4855 aagcagcaggtggaagccctctgttaacactgcatgtctgtatgtgtgtctcaagaagcca 4914
Db 4161 ACCACCCAGGTGG-GGACCCTGTGTCCCACTGCATGCTTGATGTGTGTCTCCGGAGCCCA 4219

QY 4915 cgtctagataccaagcagctctgtctctcagagga-----tcgtacacatgtgccc 4970
Db 4220 CGTCTAGGATTAAGAGCCATCTCTGCTCTCAGAGGAAGCAACCATCGTGCACATGTGCC 4279
QY 4971 ctagcaaaag 4980
Db 4280 CTAGCAAAAG 4289

RESULT 10

AF199332

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

BASE COUNT 1503 a 1166 c 1246 g 1116 t
ORIGIN

|||||
Db 1860 AAAGTTGAGAGGAAAGATGACTAATCAGGTCGACTTGTGCATTATTAACCAAGT 1919
OY 2161 aaaaaaagaagttagcgtgatactgtagacacttagccgggfgatgaagctgga 2220
Db 1920 GAAAAAGGAAGTTAGCTGATACCTTAGACACTTGAACAGGTATGAAGCTTGA 1979
OY 2221 atggaatggagagctatctgcaagagccacattggaagattacaactatctaga 2280
Db 1980 ATGGAATGGAAGGCTATTTCAGAGACCACATTGAGGAAGTTACAACTATTCTAGA 2039
OY 2281 atccaaacttgacaaccaaagtltgagctltgtltcaaggccaattggaataltccag 2340
Db 2040 ATCCAAGCTTGAAACCCAGATTGAGCTGTGTTTCAAGSCCGATTGAGATATGCTAG 2099
OY 2341 aataccgtatagcaagcagcaactgtgaatccagttcctagctcatlttgaatcaca 2400
Db 2100 AATACCTGTATGACACCTCAATGACACACTGAACTGATTCATTTGAATCTCAAA 2159
OY 2401 aatggaccgtcctctctatactccgttaccctcaatgagtcctggaatgctgaagagtg 2460
Db 2160 AATGATCTCTCCGCTATATTCGTTTACCTTCCCATGAGTCTTGCGATGCTGAGGATGT 2219
OY 2461 cccgagttcttatactgagcagcttcaataaactatggtlttgacaaggttgttcca 2520
Db 2220 CCCACAGTCTTATGTCGACAGCTTCAATTAATGTTGACAAAGTTGTGTCACCA 2279
OY 2521 gttgtagtatacaatttgggagcaagagatcctccctccagaggaagtgggagccaag 2580
Db 2280 GTTATATAGTTACAAATTTTGGACAAAGATTCCTTCCAGGGAATGGAGGCCAAG 2339
OY 2581 gaatcccttattgaagattactctctccacagatagaaggtataaataagaagaagac 2640
Db 2340 GAATCTTATGTTAAATTTACTTCTTCCAGACAGAGATGATTAACAGAGACAGAC 2399
OY 2641 aaaaacagtcaagaanaactltggaacccaatlygaacagagacttctatctcgt 2700
Db 2400 AAAACAGTCAAGAAATCTTGGAACCCAAATGGAACCAACTTTCATTTATCTCCTGT 2459
OY 2701 ccacggaagaagaatcctggaacgaatgctggaatatacccttggagtaagaagtaagt 2760
Db 2460 TCACCGAAGAGATTCGGGGAAGAAATGCTCGAAATCACCTTGGGATCAAGGTGCGGT 2519
OY 2761 tcgagaagaagaagcgaattcttagagaagatttaattgaattggaagaagcttgct 2820
Db 2520 TCGAGAGGAAGAAAGTATCTTASGAGAGATTTTATTTGAATGGAACAGTCTGTT 2579
OY 2821 agatgtagcgcgacttggatataagctgagaccatgactgtcctccatctgcaactccc 2880
Db 2580 ACATGATGAACCACTGGTACAAACTGACCAATGATGTCCTCATTTGCCACTGCC 2639
OY 2881 tcgaccttcccatatctgcccggagagcagctccatgtagagagcccaagcgaggt 2940
Db 2640 TCACCTTCTTCATATATTCACGAGACAGTGCATGAGAGAGGCCAACACCCAGGTT 2699
OY 2941 gcaaaagtgcgaagaataagtgacagtgagtgatgactgactgacagcgaagagtgagct 3000
Db 2700 GCAAAAGTGAAGAGATTAATGATACAGCAAGATATCCGACTATGACTGCGAGGCGGT 2759
OY 3001 gggagtagtgcagattatcgacacaaatggcgcgactcctcaaaagctccagctgtcgt 3060
Db 2760 GGGAGTCTGTGAGATTACGACACAGATGCGTGAATCTTCAAAAGTCCACATTAATGCT 2819
OY 3061 gccagaacaagatcgtatcaaatcatgttcaacacagaggtctccatcgataga 3120
Db 2820 GCCAACAACAGTATGCTCAACAAACCATGCTCACCATGAGGCTCTCTATTCAGATGA 2879
OY 3121 tgttatagaagaagaagtgatcgtgctgagcttagtgcctccctccctcaaaagaatgaga 3180
Db 2880 TGTATTAGGAAGACTAGTCAATGCTGCTAGTGTCTCTCTCTCAAAAGAAATGTGGA 2939
OY 3181 acaggggcaacgagagacagtgctactgagcattacaacaatltagccgaatgagtag 3240
|||||

Db 2940 ACAGGGGCTTCGAGGGACAGCTGCTACTAGGCCATTACAAATCAATTAGCCGATGACAG 2999
OY 3241 aacccgtgcatgagatgactactctctcoagtagagac----- 3280
Db 3000 ACACCGTGTATGAGACGACCATTTACTCTTTCAGAGAGACAGTCAATTTTCTACTTACC 3059
OY 3281 -----aggatltgaaag 3294
Db 3060 TCGCTCTGACACAGGCAGACCCGATGAGCATCACACAGGAGATGGAAGGATTTGTGAAGC 3119
OY 3295 agcagatagacagcaatcacaagatccagatcaacaagaacaaagcctcctctagagcg 3354
Db 3120 AGCAGTATGACAGCCATATCACAGATTCAGATCAACAGAAAGAGCGCTCTCTTAAGAGCG 3179
OY 3355 gacacacacccgtccagatccctctgaagctcctgtatataaactcataggtgtcagtc 3414
Db 3180 GACCAACACCCCTCCAGATCCCTGAACCTGTGATACAAACCTCTATGAGAGTGAATGCC 3239
OY 3415 ttcattaatgactggaagatctgcccctcctcaactgtctatcgaggtcctacccctcg 3474
Db 3240 TTCAATTAATGACTGGAAGATCTGCCCTCCTTCACTGCTTATCGAGTCTCACCTCG 3299
OY 3475 tacgggtgttccagacaagccatcaagtaactccgggaacagagagggcgca 3534
Db 3300 CACTGAGCTCTGTCCAGACAGCTCGTCAAGTACTCCCGTAAACAGAGAGAGGCGCGCA 3359
OY 3535 gcttcacagcttcccaaaaggaacatltggaaggaagtgtcatgatatagaagagag 3594
Db 3360 GCTTCTCAGCTCCCAAGGGAACACTGGAAG----- 3395
OY 3595 aaatcgccaatlgaaacttaacaatatcaaaacagtgaccgagatcagacccagatcga 3654
Db 3396 ----- 3395
OY 3655 gcaagatcacatlcgaaglatcgctcaggaatggatccacatagagggcaatcgt 3714
Db 3396 ----- 3395
OY 3715 ttcacataaactctcgagacgtgataagtgatglatctcgcggttccaagagactag 3774
Db 3396 ----- 3395
OY 3775 tgcctcgttccagcagaacagctacatgctcgltccaatcagagcgccgagaggaa 3834
Db 3396 ----- 3395
OY 3835 caggaaatcagtgcttccatcccaaaatgcaaaacagacagatgggcgtgtcgggaa 3894
Db 3396 ----- 3395
OY 3895 gaacttgacaaaagacacagcatcagltggaagacatgtgtcactgtgaagaatgagcg 3954
Db 3396 ----- 3395
OY 3955 cagcagctcgacatcgcagtgggcgccctgggttaccagltgycsaagagcgcgatcag 4014
Db 3396 ----- 3395
OY 4015 catlgyggccaaaatgtagctatgtltgltctctcagcggaaaagtcgagtgctctca 4074
Db 3396 ----- 3395
OY 4075 actcagccaaacccgaagggaggtataaaagctacagagacgtlttaagaaacacgga 4134
Db 3396 -----AGGAGGAGGTAAAAAGCTACGACGACTGTCCAGAGAAACACGGA 3440
OY 4135 gaccgggtcagcagtgagaaatggaacatgaaacccgacgaagccgaggaatccac 4194
Db 3441 GACCGGCTGACGATGAGATGAGAACTGTGATGACAAAGACAGGCCGCGGGATTCAC 3500
OY 4195 agatgcaagcatgaacagctatagctcggaaaggaatctgactctccctgggtcgcct 4254
Db 3501 GGATGGACACATGAACTGATACAGTCAAGTCAAGAAATCTGATCTTCTGCGGCTCCCT 3560
|||||

QY	4255	ggccttcgacagccagttcagtgatttccctgagtcgctggcccttcagcctagtgag	4314
Db	3561	ggccttcgacagccagttcagtgatttccctgagtcgctggcccttcagcctagtgag	3620
QY	4315	agcgcagacccctgctactcctcgaatgggtgacatcaggtgggaatgatgataaaa	4374
Db	3621	agcgcagacccctgctactcctcgaatgggtgacatcaggtgggaatgatgataaaa	3680
QY	4375	gggacagctggaggtagaatcatccgagcgccgagccttggtgtaaacaggttccaa	4434
Db	3681	gggacagctggaggtagaatcatccgagcgccgagccttggtgtaaacaggttccaa	3740
QY	4435	gacacgcccagacccgtatgtcaagtgatcgtttaagacaagagtcgtcatagccaa	4494
Db	3741	gacacgcccagacccgtatgtcaagtgatcgtttaagacaagagtcgtcatagccaa	3800
QY	4495	aaagaaaaccaaaggtggcgaagaagaccctggagccctgtacacagcctctgtcctt	4554
Db	3801	aaagaaaaccaaaggtggcgaagaagaccctggagccctgtgtatgcacagctgttcttt	3860
QY	4555	cgaaggaagaccccccaagggggggtgtaacagatcgttctggggagattatgctgat	4614
Db	3861	cgaaggaagaccccccaaggggaaggtgttacaagatcgttctggggagattatgctgat	3920
QY	4615	ggatcacaaatccttattggaggtggccagatactcttagatgaactggaaactatccaa	4674
Db	3921	ggatcacaaatccttattggaggtggccagatactcttagatgaactggaaactatccaa	3980
QY	4675	catagtgattgagtgattcaaaccttccctcctcctcctcctcctcctcctcctcctc	4734
Db	3981	catagtgattgagtgattcaaaccttccctcctcctcctcctcctcctcctcctcctc	4040
QY	4735	acccttcgaagaagaagagttcccaatcgtctcctcctcctcctcctcctcctcctc	4794
Db	4041	acccttcgaagaagaagagttcccaatcgtctcctcctcctcctcctcctcctcctc	4100
QY	4795	tcgttcataagcaataataaactgtgtgtacaaacacagcagatataaaaaacagaaga	4854
Db	4101	tcgttcataagcaataataaactgtgtgtacaaacacagcagatataaaaaacagaaga	4160
QY	4855	aacgcagaggtggagagcccttgtaacactgactgactgactgactgactgactgactg	4914
Db	4161	aacgcagaggtggagagcccttgtaacactgactgactgactgactgactgactgactg	4219
QY	4915	cgttcagagataccaagcagtcgtgtgttcacagaggaag----tcgtacacattgtgc	4970
Db	4220	cgttcagagataccaagcagtcgtgtgttcacagaggaagccatccgtgcacagtggtgc	4279
QY	4971	ctagcaagag 4980	
Db	4280	ctagcaagag 4289	
RESULT 11			
AF199325			
LOCUS	AF199325	5079 bp	mRNA
DEFINITION	Rattus norvegicus RIM2-3A (Rim2-3A) mRNA, complete cds.		
ACCESSION	AF199325		
VERSION	AF199325.1	GI:8925863	
KEYWORDS			
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 5079)		
AUTHORS	Wang, Y., Sugita, S. and Sudhof, T.C.		
TITLE	The RIM/NIM family of neuronal G2 domain proteins. INTERACTIONS		
JOURNAL	J. Biol. Chem. 275 (26), 20033-20044 (2000)		
MEDLINE	20347919		
REFERENCE	2 (bases 1 to 5079)		

AUTHORS		Wang, Y. and Sudhof, T.C.	
TITLE		Direct Submission	
JOURNAL		Submitted (27-Oct-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA	
FEATURES		Location/Qualifiers	
source		1..5079	
gene		/organism="Rattus norvegicus"	
CDS		/db_xref="taxon:10116"	
	1..5079	/gene="Rim2-3a"	
	52..4158	/gene="Rim2-3a"	
	/codon_start=1		
	/product="Rim2-3a"		
	/protein_id="AA81647.1"		
	/db_xref="GI:8925864"		
	/translation="MSAPLGRPRAPTPASOPPPPEMPDLSHLEERKTIQAVM DROKREEREOVLAKLHOPEMYKIDQVKKMGESQOQDQKGDAPTCGICHTKRFAD GGCHNSYCOTKFCARGCRVSLRSNKVMWVNCIRKQDILTKSGAMFTNSGNTQ OPDOKALRGIRSEAPOEKAKLHECTOFQPGPDSVPAVERGAGLRLQDSIKNG SGMKHOTAISMPDRKRSRPSVSRDNRKRYOSEEEREYSQVPSDSTMPSPSYADP RSOREFOYFEEDPHLNRDNRGHRHSKEYIVNDEDEVSEDEYRQREERYOARYR SDENLARYPEVPEOMRIHAEVSARHERHRSRHSVLANAELEDSRISILMDRPS RORSVERRAMENORSYMERKREAGQSSYPORTTNHSPPTPRRSPICLRDELR ADSLRKOHHLDPSSAVKRTREKMETLRDLSLSDSDSESVRPPPRHKSKGKGM OVSLSSEELASTPEYTSDDVEIESSEYKEDMEYSWLEHASMSSEASPSMLAP VPMOPSKDQRLIGRLILNKLKDGVSPPDSGAMILGLKVVGGKMTESRCAFTTKV KGLADIVGHLRSGDEVLEWNGRLDQATFEVYVNIILESEPEOVELVYVRPIGDMP RIDPSTHQAULESSSSSFESOKMDRPSISVSPMSRGMRLDVPOLSGQSSLSRT TPEVPRVOIKLMDKGVHQLVLTGLKKDLSREDGRPRNRYVYITFLPDSDKKRR TKTKVKTLIEPKMNQFTLYSPVHREPERMETLITLMDARYRESESEGLILELET ALIDDEPHWYKLOTHDVSLPLPHPSYMPRROLHGSPTRRILDRSKTISSESDV CEDGVGVSDYRHDGRDLSLTSVPQVMSNHCSPSGSHRDVIRTSWSVSP PQRNVPQGLRGPRATGHYNTISMDRHRVDHYSSESDSHFTLIPSRHQRQSEH HROGRDCEADRPYHRSRSEORPLDERTTRSSSRADTNILRSMPSMTGSRAP PSPALSRSHPTGVSQVTSPTPYTGRGRGLPOLPGKTLRGGKGLRSTVORSTE TGLAVENRMMNTRGASRESSTDSGMSNYSSEGNLFEPYRLASDSQFSPDLGLPAOL VGRQTLATPAMGDIOVGMMDKKQLEVEIIRARGLVAVPGSKTLAPVYKTYLLNGV CIARKTKVARKTLEPLYOQLLSEESPOGVLOIIVGDYGRMDKSPFMVAQILLD ELEISNNVIGMFKLEPPSLVDPLAPLTRRASOSLSLSTGSPYSRS"		
BASE COUNT		1517 a 1177 c 1257 g 1128 t	
ORIGIN			
Query Match		56.3%; Score 2802.4; DB 94; Length 5079;	
Best Local Similarity		77.2%; Pred. No. 0;	
Matches 3939; Conservative		0; Mismatches 261; Indels 899; Gaps 8;	
QY	1	gtctccctagaggttcggtcccgcaaacatgtcgttcgctcggtcgcccgccggcgccg	60
Db	21	gcttcctccttaggttcggtcccgcaaacatgtcgttcgctcggtcgcccgccggcgccg	80
QY	61	cccggtcccgcccgccggcgccctcctcaacctcctccgacgcccgaagatgcccgaactcag	120
Db	81	cccggtcccgcccgccggcgccctcctcaacctcctccgacgcccgaagatgcccgaactcag	140
QY	121	ccacctcagcgaagagagagaagaatcattcctggtctgtcattgagatcgcttaagaagaaga	180
Db	141	ccacctcagcgaagagagagagaagaatcattcctggtctgtcattgagatcgcttaagaagaaga	200
QY	181	agagagaagaagagcagtcggtgtcctcaagatcaagaagaacacacacacacacacacacac	240
Db	201	agagagaagaagagcagtcggtgtcctcaagatcaagaagaacacacacacacacacacacac	227
QY	241	gtggttcctcttagtggagctacgtgactgaataacatgcttctgagcccgacacaaa	300
Db	228	-----	227
QY	301	acaaccaatgagaagaagcccgacagacaagctgcaccacaacatttgaatgtataagga	360
Db	228	-----AAAGCTGCATCAACAATTTGAATGTATGAAGA	260

[illegible]

Db	1341	TCATATGCTCTCTTACCCCTTCGAGAGGACCCCTATACCGCTTCGATGACACCAAGCTTGAGGCG	1400
OY	1501	CGCTACCTCCCTACGAGAAACAAGCAACCACTTAGAATCCCACTCTCTGCTCTGTGAGAGAAACGAA	1560
Db	1401	TGCGCACTTCCTACGGAACCAACACCACTCTTAGATTCGCCAGCTCTGCTGTAGGAAACGAA	1460
OY	1561	GCAGAGAAAAATGygaAACCACTGtTlaagaaTgaTctctTgaTctgaTccagcattccgaTgc	1620
Db	1461	GGGAGAAAAAATGygaAACCAATGTTGAGGAATGATTTCTTGAGTTCCAGCCAGCTCTTGAGTC	1520
OY	1621	agTgagccgccccccaagcgcctataatccaaagaaagagTgtaaaTgtgccaagT	1680
Db	1521	AGTGAGGCGCGCCCAACCAAGGCTCTATTAATTCAGAAAGGAGGTAAATATGGCGCCAGGT	1580
OY	1681	ttcacttgacagcactcggagagagagcttgagcattccacacactggaTatacaagctctgtgTga	1740
Db	1581	TTTCACTGACGACACTCAGAGAGAGAGTGTGGCATTCACAGCCCTGATATATCAAGCTGTATGTA	1640
OY	1741	tgttgagcttgygaagcggagagTgtgagTgagaaagggagacagTcaaaaagggaaaaagaa	1800
Db	1641	CGTGAGATTTGAAAGGAGAGAGCGTAGTGAAGAAAGAGACA-----	1681
OY	1801	aactgagTgagcgggagTtltgtcggaTctaaacaccagTctcgagagacaaaaagaaag	1860
Db	1682	-----	1681
OY	1861	gattactatgtgtgcccactcttTgaagagagTctTggaatgtctTgagctcgaTtaa	1920
Db	1682	-----TGAGTACAGCTGCTGTGAGAGATGCGTCTTG	1712
OY	1921	ggaactcgggTgagatcactgtagTgagcaaacaccttaacgagagagatagcagaTga	1980
Db	1713	G-----CATAGCAGTAGGACATNCCCAATGTC	1739
OY	1981	taagcaacctgtgacctgtgcagccatccaaagtTgagatcgccTaatTgtgtgTatctt	2040
Db	1740	TTTTGACCCCTGTGACCTGGCAGCGCAATCAAAAGTGAAGATGCGCTTAATTTGCTGATATTT	1799
OY	2041	attaaataagcgtttbaaagatbgggagTgtTactctgagactcgaagagcaatgtcTgagctt	2100
Db	1800	ATTAAATAAAGCTTTAAAGATGAGAGTGTACTCTCGAGATTCAGAGACCAATGCTGGGCTT	1859
OY	2101	aaagTgtTgagaggaagaTgactgaaTcgaTcgaactTgtgcattTaatccaaagt	2160
Db	1860	AAAGGTTGTAGGAGGAAAGATGACTGAATCAGAGTGCAGCTTGTGCAATTTATTCACCAAGT	1919
OY	2161	aaaaaaaggaagtTtagctgtgatactgtTgaagacatctTtagccaggtTgataagtctTgga	2220
Db	1920	GAAAAAAGGAAGTTTATGCTGATACTGTAGACACTTGTAGCCAGGTGATGAAGCTTTGGA	1979
OY	2221	atgygaTggagagcTatTgcaagagagccaattTgaggaagTtlaacaatatctTgga	2280
Db	1980	ATGGAATGGAAGGCTATTTCGCAAGAGACCCACATTTGAGGAAGTTTACACACTTATTCTTACA	2039
OY	2281	atccaaacctgaaacccaagtTgagctTgtTgtTtcaaggccaattTgagatatctTgag	2340
Db	2040	ATTCCAAGCTGTAACCCACAAGTTGAGTGTGTTTTCAGAGCCGATTTGGAATATGAGCTTAG	2099
OY	2341	aatacctgatagaagcagTgcacaaactTgaaTcccaTttagctatctTgaaTctcaaaa	2400
Db	2100	AATTACTGTATACACATCAATGCAACAACTGGAATCAGATTCTAGGCAATTTGAATCTCAAAA	2159
OY	2401	aatgagcagTcctctataTatcgTtaccTcaaccatbaagTccTgggaTgtcTgagagagTgt	2460
Db	2160	AATGGAATCTTCCTCTATATTCGTTATCCCTCTCCATGAGTCTCGCATTCGCTGAGGAGATGT	2219
OY	2461	cccgagcttctatctTgagcagctTtca-----	2488
Db	2220	CCGACAGTTCTTATCTGGAACAGTTTCAAGCCAAAGCCTTAGTAGAAGAAACAACGCCCTTT	2279
OY	2489	-----ataaactatgtgtTgcaaggtTgtgtccagTttagatagTtacc	2532

Db 2280 TGTTCCTAGGGTTGAGATAAAACTATGGTTTGACAAAGGTGGTCCACGTTAATAGTTAC 2339
Oy 2533 aatttgggagcaaaagatctcccttccaaagatggaagcgcaagaatcctatgt 2592
Db 2340 AATTTTGGGAGCAAGAGATCTCCCTTCAGGGAAATGGGAGGCCAAGAAATCCTATATGT 2399
Oy 2593 taagatttactcttccaga tagaagtga laaaaaa taagaaga aaaaaa aasgtcaa 2652
Db 2400 TAAATTTACTTCTTCAGACAGAGATGATAAAAACAAGANAAACAAAAACAGTCAA 2459
Oy 2653 gaaaacttggaaeccaatgtagaacagacttattatctctgctccaggaaga 2712
Db 2460 GAAAACTTTGGAACTTCAAAATGGAACCAAGACTTTCATTTATTTCTCTGTTCCAGGAGAA 2519
Oy 2713 attccggaagaaatgtcgtgaatatacccttggatcaagcttagagtgcgaagaaga 2772
Db 2520 ATTCCGGGAACGAATGCTCGAAATCACCCCTTGGATCAAGCTCGATTCCGAGAGAGAA 2579
Oy 2773 gaagcaatcttgaagagatttaattgaattggaacaagcttgcgtagatgaagcc 2832
Db 2580 AAGTAATTCCTAGAGAGATTTTAAATGAATTGAAACAGCTCTGTGTGATGATGAACC 2639
Oy 2833 gcaatgtatagctggaagaccatgatatgtctccatgtgcacatccctccctccccc 2892
Db 2640 ACACGTGTAACAACTTCAAGACCCATGATGTCTCTCATTTGCCATCTCCCTCACCCCTCTCC 2699
Oy 2893 atactcgcgcgagcgagctccatctgagagagaccgaacgagcgcaagctgcgaagtgcga 2952
Db 2700 ATATATGCCAGGAGACAGCTGCATGAGAGAGACCACACGAGGTGCAAAAGGTGAA 2759
Oy 2953 gagaataagtacagtgaaagtgtctgactaacagctgcgaggaatgsgctggaaatgagtgc 3012
Db 2760 GAGATAAGTGCACACGGAATTCGACTATATACGTGAGAGAGGGGTGGGAATCGTGTGTC 2819
Oy 3013 agattatgacaacaatgagccgagcttccaagctccacagctgtgctggtgcgagaacaagt 3072
Db 2820 AGATTATCCGACACGATGGCGGTGATCTTCAAAAGCTCCACATTTATGTCAGAACAGT 2879
Oy 3073 catgtcatcaaatcatctgtctcaacatcaaggtctcctcatcgatgtagatgttataagaag 3132
Db 2880 CATGTCAATCAAAACCAATGCTCACACATCAGGCTCTCCTCATCGATGATGTTATAGAGAG 2939
Oy 3133 gacaaggtcatgtgtgcctgaagtgcgcctctctctcaaaagaaatgtggaagaagcgacag 3192
Db 2940 GACTAGGTATGTGTGCTTGAATGTCTCTCTCTCAAGGAATGTGGAAACAGGGGCTTCCG 2999
Oy 3193 aggaacagctgtctactgagcatatacaacaacatlaagcgaatgtagatacaacgctgtcat 3252
Db 3000 AGGGAACGCTGCTACTGGCCATTTACATACAAATTAGCGCGATGGACACACCGTGTCTAT 3059
Oy 3253 ggaatgcaactactctttagatagagac----- 3280
Db 3060 GGAGCAGCAATTACTCTTTCAGAGAGACAGTCAATTTTCTCACTTACTCGCTCTCCGACA 3119
Oy 3281 -----agggaattgtgaagcaagcagataagaca 3306
Db 3120 CAGGCGACCAAGTAGACATACCAACAGGATGGAAGGATTTGTGAAGCAGCAGATAGACA 3179
Oy 3307 gscataatcaagatccagatcaacagaacaagcgctctctctagagcggaaccacaacccg 3366
Db 3180 GCCATATTCACAGATCCAGATCAACAGAAACGGCTCTCTCTAGAGGGGACACACCCG 3239
Oy 3367 ctccagatctctggaagctctgatacaaacctcctagagtcgagtcgctctatlaagac 3426
Db 3240 CTCACATCTCTGAACTGTCTGATACAAACCTCATAGAGTCGATGCTTATTTATATGAC 3299
Oy 3427 tggaaagatctgcccctctccatcaactgtcctatcgaggtltaacacctcgtaacgggtctgt 3486
Db 3300 TGGAAATCTGCCCTCTCCCTCACCTGCTTATCGAGGCTCAACCTCGCACTGCGCTGTGT 3359
Oy 3487 ccagacaagccatcaagatctccggaagacggaagcggaagcgccagacagcttccacaact 3546
Db 3360 CCAGACCAGTCCGTCAAGTACTCCGTAACAGGAGGAAGGGGCGCGAGCTTCCCTCGACT 3419

Oy 3547 tccaccaaaaggaacatttggagagaagtgtctatgtgatatagaggagagaatctgcacaat 3606
Db 3420 CCCACCAAAAGGAACACTGGGAAG----- 3443
Oy 3607 gaacttaacaatatcaacaagctagccggaatcagaccagactbgagcaagattacca 3666
Db 3444 ----- 3443
Oy 3667 ttcgaagatcgtccagatgagatgccatatagaggggcaatactgtttccactaaatc 3726
Db 3444 ----- 3443
Oy 3727 ctggagacagtgatgtaagtgtatctgtcggttccaagactagtagtctctcgtt 3786
Db 3444 ----- 3443
Oy 3787 cagcagacaagctacatgcttcgtccaaatcagagcgccgagaggaagaacagaaataatcag 3846
Db 3444 ----- 3443
Oy 3847 tgtcttaacatccaaaatgcaaaaacagacagatggygctgctggygaagaaacttgcaca 3906
Db 3444 ----- 3443
Oy 3907 aagcacagcatcagtgagacatgtgtctcaacttggagaagaatlgacgycagccagctcga 3966
Db 3444 ----- 3443
Oy 3967 cactgcagtgagcgccctggtgtaacagtggaagaagcgcgatctagatctggggcaca 4026
Db 3444 ----- 3443
Oy 4027 aatgttagctatgtgtgtctctcaagaaaagtcgacgtgcctcctcaactcagccaac 4086
Db 3444 ----- 3443
Oy 4087 cgaagagaggtlaaaaagctacggaacgtctcagagaagcaggaagcggtgtacg 4146
Db 3444 ---ACGAGAGAGTAAAAAGCTACGACGACTGTCCAGAGACGACGAGCCGGGCTAGC 3500
Oy 4147 agtgaagatggaagactgtatgaccgcccagccagccggaatccacaagatgacagat 4206
Db 3501 AGTGAAGATGAAGAACTGGATGACAAAGACAGGCCACCGGAATCCAGGATGGACGAT 3560
Oy 4207 gaacagctatagctcggaaagaaatctgaatctccctggygtccgctgtgcctctgacag 4266
Db 3561 GAACAGTTTACAGCTCAGAAAGAAATCTGATCTTCTGGGGGTCCGCTGTGACAG 3620
Oy 4267 ccagttcagtgatctccctgtgagtgccctgtgctcagctagtggaagccagaccct 4326
Db 3621 CCAGTTCACCGATTTCTTGATGTGTGGCCCTGCTCAGCTAGTGGAGCCAGACCTT 3680
Oy 4327 ggcctactctgcaatggtgtgacatcaggtggygaatgatatgaataaaaggaacgttga 4386
Db 3681 GCGACTCTCGCAATGGGTGACATTTCAGGTGGGATGATGATGAAGGAGACAGCTGGA 3740
Oy 4387 ggtagaatcatccgggcgagcggtctgtgtgtaaaacaggttccaaagacactgcagc 4446
Db 3741 GGTAGAGATCATTCGGGCTCGAGGCTTGTAGTAAACCAAGTCCAAAGACATGCCAG 3800
Oy 4447 accgtatgtcaaggtgtatctgtttagaacaacggaagtcgtgatalagccaaaagaacaaca 4506
Db 3801 ACCATATGTCAAGGTATCTGTTAGCAACGGAGTGTGATAGCCCAAAAAGAAAACCAA 3860
Oy 4507 gttggtgagaagaagaccttgaagccctgtacacagcagctctgtccttgcaggaagccc 4566
Db 3861 GGTGGGAGAAACACGTGGAGGCCCTGTATACAGCAGCTGTATCTTCCGAAGAGAGCCC 3920
Oy 4567 ccaggggaaggtgttacaagatcatgtctggygaagatlaagtgctgatatgatacaaatc 4626
Db 3921 CCAGGGGAAGGTGTACAGATCATTTGTCTGGGAGATTATAGCCGATGATCAACAAATTC 3980

QY	4627	cttaatgaggatgagccgaagctacctaagaatgaactgaactcaacaatgagctatg	4686			
Db	3981	CTTTATGGAGATGGCCACGATACCTTTAGATGAACTGGAACTACCAACATGATTAATGG	4040			
QY	4687	atggttcaaaactctccctccctccctccctagtagatccaaactccggaactctgacaag	4746			
Db	4041	ATGGTTCAAACTCTTCCCTCCTTCTCCTCCTAGTAGATCCAACTTTGGCAGCTTGACAG	4100			
QY	4747	aagagcttcccaatcgctctctgysaaagtctcaccggaactcttactctgltcatagca	4806			
Db	4101	AAGAGCTTCCCAATCGTCTCTGAAAATTTACTAGGACCTTCTACTCTCTTCATAGCA	4160			
QY	4807	actctaaactgtgtgtacacaacaccgcgcatataaaaaacagaagaacacgacagtg	4866			
Db	4161	ACTTAAATAACTGTTGTCAACAACACGCGTTACAAAAACAGAAAGAACACACCAGGTG	4220			
QY	4867	gaagccctgtgaactgcatactgtatgtgtgtgtctacagaagcccaagctcagagata	4926			
Db	4221	G-GGACCTCGTCCCATCTGCATGCTTATGTGTGTCTCCGAGACCCACGCTTACGGATA	4279			
QY	4927	ccaagcagctctgtgtctcagagaaag---tcgtacacatgtgtgccttagcacaag	4980			
Db	4280	AAGAGCATCCTGTGCTCTCAGAGGAAGACCATCGTCGACAGTGTCCCTTACGAAG	4337			
RESULT 12	AF199328	AF199328	4851 bp	MRNA	ROD	04-JUL-2000
LOCUS						
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
FEATURES						
source						
gene						
CDS						

[illegible]

Db 681 AAGACGATTTCTATTAGAATGGCTCAGGAATGAAGCACACAGATCCAGTGCATGCC 740
Oy 841 ttcagacagaaaaagagtlccatcagtgctccagggatcaaaatcgaagatagcagcaag 900
Db 741 TTCACGACAAAAAAGAGTCCGTCAGTGCAGGGATCAAAATGAAAGATACGACCAAG 800
Oy 901 tgaagaagagagactactacaglatgltccctcagatgltacaaatgccaagatccc 960
Db 801 TGAAGAAAGAGAGANTATTACAGTATGTTCCCTTCAGATAGCAAAATGCCATGATCC 860
Oy 961 ttcgaglatgctgatatagacgattccagcgtagagcttcaatttatagaagacctggtca 1020
Db 861 ATCAGATTATGCTGATAGACATCTCAACGTGAACCTCAATTTATGAGAAACCTGATCA 920
Oy 1021 ttttaattacagagatctcacaagagagacagacatccaaagaglatatgttga 1080
Db 921 TTTAAATTATAGGGATTCTTAACAGAGAGGCCATAGACATTCCAAAGAGTATATTGAGA 980
Oy 1081 tga tgaagatgtygagagcagaga tgaalatagaagacaaagagagagagaaatacca 1140
Db 981 CGACGAAGATGTGGAGAGCAGAGATGATATGAAGACAAAGGAGAGAGAGATACCA 1040
Oy 1141 ggcacgctacagaaagtatccaatctgagcccgatcccgtaaagccaacacctaga 1200
Db 1041 GGCACGCTACAGAAGTATCCAAATTTGGCCGGTATCCGGTAAAGCCACACACTATGA 1100
Oy 1201 agaacaatcgcatccacagcttgagtggtlccagagycagacatgagagaaagcacagtga 1260
Db 1101 AGAGCAAAATGGGATTCACAGCTGAAGTGTCCCGGACAGACACGAGAGACACAGTGA 1160
Oy 1261 tgttctcttgcaaacgctgaactagaaagattccagagattctctgctgaagatgtaga 1320
Db 1161 TGTTTCTTTGGCAATGCTGAACTAGAAGATTCCAGGATTTCTGCTGAAGATGATAG 1220
Oy 1321 accatcaagcaaaagatctgtatctgaacgttagagctgcaatgagaaaccaaagatcgta 1380
Db 1221 ACCATCAAGCAAAAGATCTGTATCTGAAGCTGAGCTGCAATGGAAACACAGGATGTA 1280
Oy 1381 tccaatggaagaaactcgagagagctcagggacaaagttcttatccaaagagacctcaaa 1440
Db 1281 TTCAATGGAAGAAAGAACTCCAGAGAGCTCAGGGACAAAGTTCTTATCCAAAGGACCAAAA 1340
Oy 1441 tcaatgcttccaccctccgctcgagccctataccgcttgatagacabagactgaagcg 1500
Db 1341 TCATATGCTCTCTTACCCCTCCGAGAGACCCCTATACGCTGCATAGACCAAGAGTGAGCG 1400
Oy 1501 cgtgactccctacggaagaaacagcaccaacttagatccagctctgctgtgagaaacgaa 1560
Db 1401 TGCCGACTCCCTACGSAACAAACACACTTAGATCCAGCTCTGTGTAAAGGAAAGGAA 1460
Oy 1561 ggcgagaaaaatggaacccaatgttaaagaaatgattcttgagttcagaccagtcgagtc 1620
Db 1461 GCGAAGAAAAATGGAACAAATGTTGAGAAATGATCTTTGAGTTCAACACCAAGTCTGAGTC 1520
Oy 1621 agtggagccgcccccaaccaagagccatataatccaaagaaagagataaaatgycgcaggt 1680
Db 1521 ACTGAGCGCGCCGCCCAACCAAGGCTCATTAATCCAAAGAAAGAGAGTAAATGCGCCAGGT 1580
Oy 1681 ttcactgagcagctcgagagagagagctggtcatccacaacctgagtatcaagctgtgtga 1740
Db 1581 TTCACTGACACAGCTCTGAGAGAGAGATTGGCATCTCACGCTGAGATATCAAGCTGTGATGA 1640
Oy 1741 tgtgagctcggaagcgagagtgtagtgaagagagagagacagttcaaaagggaaaaagaa 1800
Db 1641 CGTGGAGATTGAAAGCGAGAGCTAGTGTGAAGAAAGGAGACAT----- 1682
Oy 1801 aactagtgagcagagagtttgtgcgattctaaaccaaggtctgagagacaaagaaag 1860
Db 1683 ----- 1682
Oy 1861 gatgtactatgtgtgccactcttgyaagaggtttggaatggtctgagctcagattaa 1920
Db 1683 -----GGAATACAGCTGTGTTGAGAGCATGCGCTTTG 1712

Oy 1921 ggaactctgggttagatactgtagtagcacacccttaacagagagatagccatagta 1980
Db 1713 G-----CATAGCATGAGGACATCCCAATGTC 1739
Oy 1981 taagcacctgtgactgtgcaagcatccaaaga tggagatcgactaattggtcgalttt 2040
Db 1740 TTTGCACCTGTGACTGTGAGGCAATCCAAAGATGGAGATCGCCTAATTGGTGTATTTT 1799
Oy 2041 atbaaataagcgtttaaaga tggaggtgtactcgagacttcggaaacatgctgggctt 2100
Db 1800 ATTAATAAGCCGTTTAAAGATGGAGCTGATCCGAGATTCAAGACCAATGCTGGGCTT 1859
Oy 2101 aaaggttggagggaaaga tgaactgaatcaggtcgacttgtgcatlataccaaagt 2160
Db 1860 AAGGTTGTAGAGGAAGAGATGACTGATCAGTGCAGCTTGTGCATTTATTACCAAGT 1919
Oy 2161 aaaaaaagaaagtttagctgatactgtagaaca tcttagaccaggtgataagcttgya 2220
Db 1920 GAAAAAAGAAAGTTTACTGTATCTGTAGGACATCTTAGACAGGTGATGAAGTCTTGA 1979
Oy 2221 atggaaatggaggtatctgcaagagccaaattggagaaattacaacatlatctaga 2280
Db 1980 ATGGAATGAAAGGCTATTTGCAAGAGCCACATTTGAGGAATTTNCAACATTATTCTAGA 2039
Oy 2281 atccaaacctgaaccacaagltgagctgltgtlccaaagccaattggagabaltcctag 2340
Db 2040 ATCCAAACCTGAAACACACAGTTGAGCTTGTTTCAAGGCCGATTTGAGATATGCTTAG 2099
Oy 2341 aatacctgatagaacgcatgcacaaactggaalccaagtcttgatattgaaatccaaaa 2400
Db 2100 AATACCTGATAGCACTCATGACAACTGGAATCCAAGTCTAGCTCATTTGAATCTCAAAA 2159
Oy 2401 aatggacgttccctctatccgttaacctcaaccctagatccctgagcatgtagagatgt 2460
Db 2160 AATGGATGCTCCGTATATGCTTACTCTCCATGATGCTCGCATGAGGAGATG 2219
Oy 2461 cccgagttctlatctgagacagcttccaataaactatggtttgcaaggtgtgtcacc 2520
Db 2220 CCCACAGTTCTTACTGTGACAGCTTTCAATAAACTATGTTTGAACAGTTGTGCACCA 2279
Oy 2521 gttgatagttaacttltggagacaaagatctccctccagggagatgtagagcgcaag 2580
Db 2280 GTTATATGATTCAATTTTGGAGCAAAAGATGCTCCCTTCCAGGAAAGATGGAGGCGCAAG 2339
Oy 2581 gaatccatgtttaagatttacttccctccagatagaaagtgataaaataagagaagaac 2640
Db 2340 GAATCCTTATGTTAAATTTTACTTCTTCCAGACAGAAAGTGAATAAAMCAAGAGAAAC 2399
Oy 2641 aaaaacagtcaagaanaacttggaaacccaatggaaacagacttcaattatctcctgt 2700
Db 2400 AAAAACAGTCAAGAAACCTTTGAAACCCAAATGGAACCAAGACTTTTCAATTATCTCCTGT 2459
Oy 2701 ccaacgagagagatctcgtagaagaaatgctggaatlaaccccttggatcaagctagat 2760
Db 2460 TCACCGAAGAGAAATCCGGGAAAGAAATGCTGAAATCACCCCTTTGGATCAACATCTGAGT 2519
Oy 2761 tgcagaaagaagagcgaattctcttagagagatlttaattgaatttggaaacagcttgc 2820
Db 2520 TCGAGAGGAAGAAAGTGAATTTCTTAGAGAGATTTTAATTTGAATTTGAAACACACTCTGTT 2579
Oy 2821 agatgatgagcgagctgtgataagctgagaccaca tgaatgtctcctcatgtcacctcc 2880
Db 2580 AGATGATGAACACACCTGTTACAACTTCAAGCCATGATGATGTCCTCATTTGGCACATCCC 2639
Oy 2881 tgcgcttcccatatctgcccggagagcagctccatlgagagagcccaagcgcgagct 2940
Db 2640 TCACCTTTTCCATATATGCGACGAGACAGCTGATGAGAGAGCCCAACACGCAAGTT 2699
Oy 2941 gcaaaagtcgaagaagaatagtgacagtgaaagtgctgataagactgagagatggcgt 3000
Db 2700 GCAAAGGCTGGAAGAGAAATAGTGACAGCGAAAGTATCCGACTATGACTGCGAGAGCGGCGT 2759

[illegible]

Accession	Position	Sequence	Length
Dp	3216	-----AGGAGAGGTAATAAGCTTAGCAGCATCTCTCCAGAAACACGAGACCGG	3266
Oy	4141	gctagcaagtgaagatgaagaactgataaccgcgaagccagccgaagatccacacatg	4200
Dp	3267	GCTAGCAAGTGAAGATGAGAGAACTGGATACAAAGACAGCCGAGAAATCACAGGATGG	3326
Oy	4201	cagcatgaacagctatagctctcgaaagaatctgactctccctgggtgcgcctgc	4260
Dp	3327	CAGCATGAACAGTTAACACTCAGAAAGAAATCTGATCTTCTCGGGGTCGCCCTGGCCTC	3386
Oy	4261	tgacaagccagttcaagtgttccctggaatggcctggcctgctcaagctaatggagcgca	4320
Dp	3387	TGACAGCCAGTTTCAACGCAATTTCTTGATGTCGTGGCCCTCTCCTCAGCTAATGGACGCA	3446
Oy	4321	gacctgtctactccctgtaaatcgggtgtaacatccagttggaatgataaagaagaca	4380
Dp	3447	GACCTGGCCGACTCTCTGCAATAGGCTGATTCAGTGGGGATGATGATTAAGAAAGGACA	3506
Oy	4381	gctggaagtagaanaatcactcgcgycgcgcgcctctgtgtgtaaaaccaggtctccaaacact	4440
Dp	3507	GCTGAGAGTAAAGATATCATCCGGGCTCGAGGCTTGTATTAACACAGGTTCCAAAGACT	3566
Oy	4441	gccagcacccgtatgtcgaagtgtatctgtttgacaacggagatctgcatagccaaaagaa	4500
Dp	3567	GCCAGCACCAATATGTCAGAGGTGTACTCTTTAGACAAACGGAATCTGCATATGACCAAAAAGAA	3626
Oy	4501	aaccagaagtgcgaaagaagacccttggaagccctctgacacagactctgtctcttcgaaga	4560
Dp	3627	AACCAAGTGGCGAAGAAAGACGCTGGAGCCCTGTATACAGACGCTGTATCTTTTCGAAGA	3686
Oy	4561	gagcccccagaaggagagtgcttacagatcatgtctctgagagatlaagtctgatgatca	4620
Dp	3687	GAGCCCCCAGGGGGAAGGTGTACAGATCATTTGTGGGGAGATTATGGCCGATGATCA	3746
Oy	4621	caaatcctttatggaagtgtgccacagatacctcttagatgaactggaactaccacaagt	4680
Dp	3747	CAAAATCCCTTATGGGAGTGGCCAGATCACTTATGATGAACATGGAATTCACATCAACATGCT	3806
Oy	4681	gattggaatggttcaaacctctccctccctccctccctccctccctccctccctccctccct	4740
Dp	3807	AATTGGATGGTTCAAACTCTTCCCTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3866
Oy	4741	gacagaagaagcttcccatcgtctctgaaagtctctacccgaacctcttactctcgttc	4800
Dp	3867	GACAAAGAAGACTTCCCAATGCTGTCTGGAAGTCTACTGAGACTTCTTACTCTCGTTC	3926
Oy	4801	atagaacaataaactgtgtgtcacacaacacagcagatatacaaaaacagaaagaaacgca	4860
Dp	3927	ATAGCAACTATTAATAAAGTGTGTCAACAAACACAGCGTGTACAAAACCGAAGAAACCCACC	3986
Oy	4861	caggtggaagagcccttggaacactgcatcgtctgtaagtgtgtcttacaagaagccaaagtcta	4920
Dp	3987	CAGGTGGG-GGAGCCCTGGTCCACATGCAATGATGTGATGTGTGTCTCCGAGCCCAAGTCTA	4045
Oy	4921	gggataccaacagcagtcctcgtgtcttccaaaggaag----tcgtacacatgtgtgccctagca	4976
Dp	4046	GGGATTAAGAGACCAATCGTGTCTCTCAAGAGGAGCCATCCGTCGACATGTGTGCCCTTACA	4105
Oy	4977	aaag 4980	
Dp	4106	AAGG 4109	

RESULT 13

ABO18294 3854 bp mRNA PRI 16-JUN-1999

LOCUS ABO18294 Homo sapiens mRNA for KIA00751 protein, complete cds.

DEFINITION ABO18294

ACCESSION ABO18294.1 GI:3882222

KEYWORDS

SOURCE Homo sapiens adult male brain cDNA to mRNA, clone_1b:pluescriptII

SK plus clone:hh04424.

Db	1569	AAGCGCTAAAGATGGAAGTGTACTCGAGATTTCAGAGACAATGCTTGGCTTAAGTT	1628
Qy	2108	gtagagagaagaatgactgaatcaggctgcagcttgcattattaccagaataaaaa	2167
Db	1629	GTATGAGGAAAGATGACTGAATCAGGTGGCTTGTGCATTTATTACTAAAGTAAAAA	1688
Qy	2168	ggaatttgctgatactctgtagacactttagaccaggtgatagaagctcttggaatgaat	2227
Db	1689	GGAAGTTTAGCTGATATCTATGAGACATCTTAGACACAGTATGAAGTATGAAGTGAAT	1748
Qy	2228	ggagagctatctcaagagaccacatttgagaagttacaacattaccattaccagaatccaa	2287
Db	1749	GGAAGACTCTCGCAAGAGACCCACATTTGAGAGAGTGTACACATCATTTCTAGAAATCCAA	1808
Qy	2288	ccctgaaccaagaattgagctgtgttctcaagccaattgagatalctcctagaatacct	2347
Db	1809	CGTGAACCCAGATAGAACTTGTAGTTTCAAGGCCCTATTGAGATATACCGCAATACCT	1868
Qy	2348	gatagcacgcatagcacaacttggaaatccagttctagctcatttgaatctcaaaaaatggac	2407
Db	1869	GATAGACACATGTACACACTGGAGTCCAGTCTAGCTCCTTTGAATCTCAAAAAATGGAT	1928
Qy	2408	cgctcctctatatcgttaacctcaacctagtcctgagcagtcgtgaggaatgtcccgag	2467
Db	1929	CGTCTCTATTTCTGTACCTCTCCCATGAGTCTGGAATGTTGAGAGGATGTCCACAG	1988
Qy	2468	ttcttatctgcagacgtcttcaataaaactatggttgcagaagttggtcaccagttgata	2527
Db	1989	TTCCTATTCAGGCAACTTTCATTAATAAACTATGGTTGACAAAGTTGGTCAACCAATTAATA	2048
Qy	2528	gttaaatcttggagcaagaatgtccctctcagggagaatgaggagccaagaatccct	2587
Db	2049	GTTACATTTTGGGACCAAAAGATCTCCCTTCAGGGAAATGGAGGCCAAGGAATCCT	2108
Qy	2588	tatgttaaatattactctctccagataagaatgataaataagaagaagaacaaaaa	2647
Db	2109	TATGTTAATAATTACTTCTTCTCCAGACAGAGATGATTAATAACAGAAAGACTAAAAACA	2168
Qy	2648	gtcaagaanaacttggaaacccaaatggaacagacttcatlctcctgtccacoga	2707
Db	2169	GTTAAAGAAAACATTTGGAAACCAATGGAACCAAAACATTTATTTCTCCAGTCCACCGA	2228
Qy	2708	agaagaattccgtgaacgaatgtcgaaataacccttggataagctagaagtcgagaa	2767
Db	2229	AAGAAATTTCCGGGAACGAAGCTTAGAGATTACCTTTGGATCAAGCTCGGTTCCGAGAG	2288
Qy	2768	gaagaagcgaaattcttgaagagatlttaattgaattggaacagcttgcagatgat	2827
Db	2289	GAAAGAAAGGAATTTCTTAGCGCGATTTTATTTGAATTGAACAGCATTTATAGATGAT	2348
Qy	2828	gagcgacactggtataagctgcagaccacatgatgtcctcctcattgcacccctcgccct	2887
Db	2349	GAGCCACATTTGGTACAAACTTCAGACGATGATGTCTTCTCATTTGCCACTTCCACCCT	2408
Qy	2888	tcccatattctcccccggaggcgctccattgagagagcccaacgcgcgaagctgcagaag	2947
Db	2409	TCTCCATATATGCGACGAAGACAGCTCCATGAGAGAGCCCAACGAGGTTGCAAAAG	2468
Qy	2948	tcgaagaagaatagtcagagtgctgcatacagctgcagagtgaggaatggcgttggagta	3007
Db	2469	TCAAAAGAGAAATAGTATAGTGAAGTCTGTGACTATGACTGTGATGATGGAATTTGCTGA	2528
Qy	3008	gtgtcaagatlatcgacaaatggcgcgactcttcaagctccacgltgtcgttgcagaa	3067
Db	2529	GTATGAGATATTCGACATATGCTGAGATCTTCAAAAGCTCAACATATATAGTCCAGAA	2588
Qy	3068	caagcatctcatcaaatatgctacacatcaaggtctcctcatcgaatgagatgtata	3127
Db	2589	CANGTATATTCATCAAAACCACTGTTCAACCTCAGGGGCTCTCTCATCAGATGAGATGATA	2648
Qy	3128	ggaagagacaaggtcatgtgtcgcttagtgcctccctcctcctcaaa---ggaatgtgaaacag	3184

Db	2649	GGAAGACTAGATCATGTGTCACCCAGTGTCCCTCCACAAAGTCCGAATGTGAACAG	2708
Qy	3185	gggcaaccgaggggaacgylct---actgccaattacaacaacatlagccgaatgataga	3241
Db	2709	GGGCTTCGAGGGACCCGCACTATGACCGGACATTAATTAATCAATTACCCGATGAGAGA	2768
Qy	3242	caacgtatcatgtatgataccactctcttcaatagatagagacgggaatttgaagcagagat	3301
Db	2769	CATCGTATGGAATGATCACTTAATTTCTCCAGATTAAGACAGGATTTGTGAAGCAGAT	2828
Qy	3302	agacagccaatcacagatccagatccaacagaacagcgctcctccagagcgagaccac	3361
Db	2829	AGACAGCCATATCAGATGCCATCATCAACAGAAACAGGCGCTCTCCTTAGAGCGGACCAAC	2888
Qy	3362	accgcgtccagatcctctgaacgltccgtatatacaaacctcatlaggtcatgtcctcatla	3421
Db	2889	ACCCGCTCCAGATCCACTGAACGCTCTGATACAAACCTCATGAGTGTGATGCCCTTCATTA	2948
Qy	3422	atgactgggaatcttgccctcctctcaacctgactcttaccaggtctcaacctcgtacccgg	3481
Db	2949	ATGACTGGAAGATCTGCCCTCTCTACCTGCTTATTCGAGGTCTCATCTGTACTGGG	3008
Qy	3482	tcgtccagacaagcccatcaagtaactccgaggaacagagcaaggggcgcagacatcca	3541
Db	3009	TCTGTCCACACAAGCCCATCACTACTCCATGCGGAGAGAGAGGCGCCACACTTCCA	3068
Qy	3542	csagctccaacaaagggaacatltgagagaaatgtctatgtatataagagagaatcgc	3601
Db	3069	CAGCTTCCACCAAGGAGAACGTTGGATAGAA-----	3099
Qy	3602	caaatgaacttaacaataacaacaggtagccggatcgaaccccaactcgtggaacagat	3661
Db	3100	-----	3099
Qy	3662	taacatcgaagtatcgtccagatggtatccacatagaggggcagatcttccact	3721
Db	3100	-----	3099
Qy	3722	aaatcctcggaagtgatgtaagtgatgtatctgcglttcaagactagtagtctct	3781
Db	3100	-----	3099
Qy	3782	cglttcagagcaacaagctcaatgtcgttccaatcagagcgccgagagaagaacagaaa	3841
Db	3100	-----	3099
Qy	3842	atcagtgcttlatcatccaatgtcaaaacagacagatggtgcgtgtcgtggagaacttg	3901
Db	3100	-----	3099
Qy	3902	accaaaagcacaagcatcagltgagagacatgtctcaactggaagaatgaagcgagccag	3961
Db	3100	-----	3099
Qy	3962	tccgacactgcagltgggcgcctcgtgtaccagtggaagaagcggcgactatagatggg	4021
Db	3100	-----	3099
Qy	4022	gccaaatgtagctatgtgtgtctctcaagaaaaagtcgagtgctcctcaactcagc	4081
Db	3100	-----	3099
Qy	4082	caaacggaaggaggtataaaagctacggagcagtgltcagagaagcacggagaccggg	4141
Db	3100	-----AAGCAGGAGGTAAAAAACTAAGAGCAGCTGTCCAAAGAAATTAAGAAACAGGC	3152
Qy	4142	ctgacagtggaatgagagactgtagaccgcgcagccgagccggggaatccaaagatggc	4201
Db	3153	CTGGCGTGGAAATGAGATGATGATCTGACAGGCAAGCCGAGGTCTACAGATGAT	3212
Qy	4202	agcatgaacagatlatagctcgcgaaggaatctgatctctcccttgggtcgcgcctgacct	4261
Db	3213	ACCATGAACAGCTACAGCTCAGAGAAGAAATCTGATTTTCCCTGTGTCTGCTTGCCCTCT	3272

OY	4262	gacagccagcttcaagtgatcttccttgatggccttggccctgctcaagctagtgaggacgac	4321
Db	3273	GATAGCCAGTTTCAGTGATTTTCCTGGATGGCCTTG6CCCTGCTCACTAGTG6GACGCCAG	3332
OY	4322	acccttgactactccttgcgaatgggtgacatcaggttgggaatgatgatgaataaaggagac	4381
Db	3333	ACTCTGCGAACACCTGCATGCAATGGGTGACATTCAGGTAGGAATGATGGACAAAAGGACAG	3392
OY	4382	ctggagtgatgaatacatccggcgcgccgtctggtgtaaaacaggttccaagacactg	4441
Db	3393	CTGGAGGTGGAATTCATCCGGGCCCGCGGCTTGTTGTAAMACCAGTTCACAGCACCTG	3452
OY	4442	ccagacccgatgtgcaagtgatgtatctgttgtaacaaggagtctgcatagccccaaaagaa	4501
Db	3453	CCAGCACCGTATGTAAAGTGATCTATTATGATTAACGGAGTCTGCATATGCCCCAAAAGAAA	3512
OY	4502	accaaagtgcggaagaagaccctcggaagccctgtacacagcagctctgtctctgaagag	4561
Db	3513	ACAAAAGTGGCAGAGAAAACGCTGGAAACCCCTTTACAGACGCTATATCTTTGCAAAAG	3572
OY	4562	agcccccaggggagggtgtgtacagatcatgtctggtggagattatggtcgtatgatac	4621
Db	3573	ATCTCCCAAGAGAAAATTTTACAGATCATGCTGCTGGGAGATATATGGCCGATGGATCAC	3632
OY	4622	aatcctttatgtggaatgtgcccagatactcttgatgaactcggagacatcatccaatggtg	4681
Db	3633	AAATCTTTATGGGAGATGGCCCAAGATATCTTTAGATTAACATGAGAGCTATCCAAATATGGTG	3692
OY	4682	attgtgagttcgaactcttcctcctcctcctccctagtagatccaactcggagaccttg	4741
Db	3693	ATTGATGATGTTCAAACTTTTCCACACTTCTCTCCCTAGTAGATCCAACTTG6CTCTGTG	3752
OY	4742	acaagaagaagctccccaatcgtctctgtgaaagtctacacggacctcttactcgttca	4801
Db	3753	ACAAGAAGAGCTTCCCAATCATCTCTGGAAAGTTCAACACTGGACCTTGTACTTCGTTC	3812
OY	4802	tagcaact---ataaactgtgtgtcacacaacaccagcgatac	4840
Db	3813	TAGCAGCTGTAAAAAATTTGTGTCTACAGACACCGAGCTTAC	3854
RESULT 14			
AF007836	LOCUS	5655 bp mRNA	15-AUG-1997
DEFINITION	Rattus norvegicus rab3 effector (RIM) mRNA, alternatively spliced, complete cds.		
ACCESSION	AF007836		
VERSION	AF007836.1 GI:2317777		
KEYWORDS	.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 5655)		
TITLE	Wang,Y., Okamoto,M., Schmitz,F., Hofmann,K. and Sudhof,T.C. Rim is a putative Rab3 effector in regulating synaptic-Vesicle fusion		
JOURNAL	Nature 388 (6642), 593-598 (1997)		
MEDLINE	97394473		
REFERENCE	2 (bases 1 to 5655)		
AUTHORS	Wang,Y. and Sudhof,T.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUN-1997) Molecular Genetics, UT Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA		
FEATURES	Location/Qualifiers		
Source	1..5655		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	/tissue_type="brain"		
gene	1..5655		
	/gene="RIM"		

[illegible]

OY	475	tttcattatgccaacaacaaagttctctgtcctcgaatgtagagtcgaagtctcttatacgtccaa	534
Db	879	CTCCATTATGTCACACCAAACTTCTGTGTCACAGCGCCGAGGCCCGTGTCTCTCTGGATTCGAA	938
OY	535	caa-----gattatgtgtgtgtatattgttgcggaacaaagaataatctc	582
Db	939	CANATGAGCAAAAGTGGTTATGTGGGTATGCAAATTATGTGCAAAACCAACAAATAATTCTT	988
OY	563	caataatcagaagcatgtgtttataatagtgtgtctaaacacatgcagcaacctgatca	642
Db	999	AACGAAATCTGGAGCGTGGTCTCTTGGAAAGTGGCCCT-----CAGCAACCTTGTGTCA	1049
OY	643	aaagttcctc- gaaggtcttgaaatgaaagaagcccttaagsgaagaagaacaaactac	701
Db	1050	AAATGGAGATCTGAGTGACACGCGCCACAGAGTGCTGTGATCTGAGTGTCCTCAAGACAAAGAA	1109
OY	702	acgagcagcccccagattccaagaagagcccccagtgbaactatcagtaactgcagtttgagaag	761
Db	1110	ACCAAGGCTCCAAAGAGCGATCAAAAGTCTCAAGAGCCCTTGATGATACAGACGTCTCTCTTC	1169
OY	762	gccgagctcatatgtgtctacaagaacagatatactattaaatagtacaagaagtgaagcacc	821
Db	1170	CCAAAGACATGCTACTACCCCGCGTGACACCGTGCACAGAAACMAAGGGCGTGAAGCCCTCACA	1229
OY	822	agattgcagtgacatgtccttcaagacagaagaagaatccatcagttgtccagagatcaaa	881
Db	1230	GCAAGCCTTGGGTCTCCGTAACACAGACGAGCGATCAAAATCAAGAAAGCAGCCACCGAGGGA	1289
OY	882	atcgaagatatacagcaaaatgaaagaagaagatactatacagatgtgtccttcaagt	941
Db	1290	AAGGAGAAAGAGGCTCCAGGGCTTTCACAG-----CAGATGCGCAAGGAGGACCGACAGAG	1341
OY	942	gtacatgtcccaagatctcctcttgatgtatgtcgtatagatagacatctcaagcgtagccat	1001
Db	1342	ACGAGCGCAAAAGTCTCCCAAGTGTGTGTCAACCCGGGCAAGGATCGCGGATGAG	1401
OY	1002	tttatgaagaacctgtgtcatttaattacaaggaattctaacaggaagagccatagaacat	1061
Db	1402	ACGGAGAGAGAAAGAGAGGGGGGGAACCCGCGATTGGAGAAAGGGCCCTCCACAGGATAC	1461
OY	1062	ccaaagagatatttgtgtatgtatgaagaatgttgaagcgaaagtatgaatataatgaagaacaa	1121
Db	1462	TCAGAACCGGGCT-----GAGAAACCGCAATGTGCAGAGGTGTGGGAGAACACAGAGACAGA	1516
OY	1122	ggagagaggagaataaccgagcagctacagaagtataccaatacttgcgcggtatccg	1181
Db	1517	GGAAGGAGGAGGAGTACAGACTAGGTACCGCAGCGACCTTAACCTGCGTCCGTACCCGG	1576
OY	1182	taaagccacaaccttaagaagaacaaatgtcgcatccaagcttgaagtgttccaaggcacgac	1241
Db	1577	TCGAAGCGCGCGCAGAGGAGCAGCAATGTGCGTATGCAAGCCCGCGGTGTCCCGAGCGAGCG	1636
OY	1242	atgagaggaagacagatgtatgttcttcttggcaaacgttgaactcagaagaattcagaattt	1301
Db	1637	ACGAGCGGGCCACACCGACGATGCGCCTCCCGCACACCGAGGACGCTGCCGCGCGCGG	1686
OY	1302	ctctcttaagagatgataagaacatacaagacaaagatctgtatcttgaacgttagagctgcaa	1361
Db	1697	CTGAGGCCACGGGGGCGCAAGCGCGCGGCCCA-----CGGCCA	1735
OY	1362	tggaaaaaccaaagatcgttatcttaatgtgaaagaactcgaagaagctcagagaacaaagtctc	1421
Db	1736	GGGTCTCTCCCCGAGTCCCGCGCGCAAGCCCGGGCGCGGCCACACCTC-----	1786
OY	1422	atcccaaaagacccaataatcattagtccttccaaacccctgcgaggaagccctatacagcttg	1481
Db	1787	-----CCACCGAGACAGGGGCCAACCGCGCGCGGCCACGCCCGGATCCGCGAG	1834
OY	1482	atagaccagacatgaagcgcgctgtactccctacagaaacgaacacactatgatccagct	1541
Db	1835	AGCCACCCAGAGCGCGCTCTCCCGACCGCTCTCGTTAAGCAGGGCCCGGCTGTGACCCGGGCTT	1894

QY	1542	ctgc-----gtgagggaaaacgaaagcgagagaaaaatggaaccatgftlaagaaatgatt	1595
Db	1895	CGGCGGCTCTTCTGGGCAAGGCGCGGAAAGCGGAAGAGATGCTGGCGAAGCACT	1954
QY	1556	ctttagttcaagaccagaccgagtcagtgagcgcccccaccagaagcctcataatcca	1655
Db	1955	CGCTAAGCTCCCAATAGTCCGAGTCCGTCGGGCCATTCGCCCCCAAGCTTACCGGCCCA	2014
QY	1656	agaaagaggtlaaaatgycgcacagtttcactagtcagcctcgagagagagctgycatcca	1715
Db	2015	AGCGGGAGAGCGCAAGAGACGTCAGATCGTGTCAGACCTCGGAGGAGGAGCGGTCTCA	2074
QY	1716	caacctgatatcaagcctggtgatgtgagcctgagaaagcgagagtggtgagtgaaag	1775
Db	2075	CACCGGAGTACACGAGCTCGCGAGGACGTGGAGCTTGAAGAGGAGAGCCTTGAGCCGAGAAG	2134
QY	1776	ggagcagtcataaaagggaaaaaaactagtgagcagggagtttgcgatatctaaca	1835
Db	2135	GTGAC-----	2139
QY	1836	ccaagtlctgagagacaaaaaagatgtactatgctgycacactcttggaaagagatt	1895
Db	2140	-----TTGGATTT	2146
QY	1896	tggaaatggtctgagcctaagatlaaagacctgvggtgatatacctgtatgacacaacc	1955
Db	2147	ACTACTGTTGGATCCCGCAGCTGGCACAGCAGGGAACGTCGCCCTTATCAG-----	2198
QY	1956	ttaagagagagatagccaatgtaaaagcaccctgtacccgtagagcatccaagaatg	2015
Db	2199	-----TTGGCATCTGTTACGTGGGAGGCCGCTTAAAGAGG	2233
QY	2016	gagactcgccctaattgctcgtatttataataaagacgtttaaagaatgagatgtaacc	2075
Db	2234	GAGATCGACTAATTCGGCGGTGTTATTCTTAAACAAAAGAACAAACA-----TGCCCA	2284
QY	2076	gagactcagagacaaatcctgggtctaaaggttgtgagaggaagaatgactgaatgaatc	2135
Db	2285	AAGAATTCAGAGTCACTTATTATGCGTGTGAAGGTGGTTGGAGGAAAATGTAGCGACTTAAAGGC	2344
QY	2136	gaccttgacattatccccaagtaaaaaaagagaatttagctgtaactgtlagaacatc	2195
Db	2345	GCGTTGGTGTTCATCACCAAGATGAAGAGAGGAGCGCTGACAGACTCGTGGACACC	2404
QY	2196	ttagaccagtgatataagtccttggaaatggaatcgagagcctatctgaagagagccaatltg	2255
Db	2405	TAAAGAGACAGGGACGAAGTCTTAGAGTGAATGATTAACCCCTGCCGGGAGCAACAAAG	2464
QY	2256	aggaagtttaacaacatlatctagaaatccaaacctgaaccacaagtltgaagctgtgttct	2315
Db	2465	AAGAAGTTTACAACTTATTCTTAGAATCAAAATCAGAACTTCACCTTGATGAGATTATTGTTT	2524
QY	2316	caaggccaattggagatatctcctagaatacctgatatagcaagatgacacatcggatacca	2375
Db	2525	CAAGGCTATTGTGTGATCATCCCCAGAGATCCCTGGAAGTTCCTCCCTCGGAGTCTCA	2584
QY	2376	gttcttagcctatgtaactcctoaanaaalgagacgctccctctatctacgttlaactaca	2435
Db	2585	GTTCAGATTCTCTTGAAATCTCGAAAATGGAAGGCGCTTTCTATTCTGTATTATTTCTCCAA	2644
QY	2436	tgaatcctggaatgctgagaggaatgctccgcagttctctatctlgagacagcttccaataaac	2495
Db	2645	CCAGGCTTGAGCTCTGAAAGATGCCGCCCAAGTCTTTACCAAGGCAACTCTCAGTGAAGC	2704
QY	2496	tatggtttgacaaggttgttccacagtttgatagttaacaattltgggaggaagagatccc	2555
Db	2705	TATGTATATTAAGTGGGCGACACGATGATGTAAATGTTTTCACACCAACAGATCTAC	2764
QY	2556	cttccagggagaatggaagcgcaagaatccttatgttataagatttactcctccagaata	2615
Db	2765	CCCTTAGATTAATGCGCGTCCCAAGGAATCCCTATGTATAAAATGTATTTTCTTCCAGATA	2824
QY	2616	gaagatgataaaaataagagagaaacaaaacagltcaagaaaaactltggaaaccaaatgga	2675

Db 2825 GAAGCGACAAAGTAAAGAGAAACCAAAACAGTAAAGAACTTCTAGAGCCAAATGGA 2884
QY 2676 accagacttcatltaletcctcgtccaccggaagaattccgtgaaacgaatgctgaaa 2735
Db 2885 ACCAGACATTGTTGCTACTACAGCTACATGTAAGATTTTCGAGACGAATGTTAGGA 2944
QY 2736 ttacccttggatcaagctagagtlccgaaagaagaagacgaattccttaagagaattt 2795
Db 2945 TTACCGTGTGGGACCGAGAGATGACAGACGAAGAGATTTCTTGGAGAGATCC 3004
QY 2796 ttattgaattgaaaagcttctgctagatgatagccgactggtatagaagtcgaacc 2855
Db 3005 TCAATGAGTTGGAAACAGGCTTTTATGATGAGCCCATTTGGTATTAACCTCCAGACAC 3064
QY 2856 atgatgtccctaatgtgcacatccctccgacctcccatatctgccccggagagcgtcc 2915
Db 3065 ATGACGAATCTTCACTACCTCTGCTCAGCCATCACCGCTTCATGCCAGGCGCATATTC 3124
QY 2916 atggagaagcccaacgcgagagctgcaaaagtccgaagaataatgtacagtgaaatgt 2975
Db 3125 ATGGAGAGAGCTCCACAAAAGCTACAAAGATCTCAGCGAATCAGTATGATGACATCT 3184
QY 2976 ctgacacagactgcgagagatgagcgctgagagtaatgtc-----agattatcgacaaatg 3029
Db 3185 CAGATTATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3244
QY 3030 gcgcgacttccaagctccacgctgtcgtgccaagacaaagtcaltgcatcaaatcatt 3089
Db 3245 CTAGAGAGAGTAAAGAACCAACGTTAACGATGCCAGACAAAGAACTACATCACCC 3304
QY 3090 gctacacatcaaggtctctcctcctcagtagatgttatagaaggaacaaagtcatgtgtcgc 3149
Db 3305 GCTCACGTTCCGCTGCTCCTCATCGCGCATGATCAGGGAGGCTCGTTCACGTTTAC 3364
QY 3150 ctgagtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3206
Db 3365 CAAATGTCCTTACAGAGAGAGCTTAAATTAATTCACACAGAAAGTGCACGTTTCTC 3424
QY 3207 ctggtccattacaacacaaatgacgaa-----tggatagacacogtltcatgtatg 3257
Db 3425 CAACCCGACACCATGATGCTCCGGAAGCCCGCGCATCACAGATCCAGACATGCGAAA 3484
QY 3258 accacactcctttagatagagacagga-----cttggaaagcagcagata 3302
Db 3485 GTCAAATTTCTGAGAGCCAGCAGATGAGCTTCTCATGCTGCCACAGACAAACAGGAC 3544
QY 3303 gacagcatatcagatccagatccaagaacaagcgctcctcctcctcctcctcctcctcctc 3362
Db 3545 GAAGTGCAGAAAGCTTACATGACATGACAGCTGACAGCCCTCTCTTGACAGAGGTTAGGA 3604
QY 3363 cccgctccagatcctcctgaaagtcctgatatcaaaactcctaagtgagtcg-----atgc 3413
Db 3605 GTGCTAGTACCAACTGCTTGAAGCCAGATACTAGTTTCATTCCACGAAGAGAAAGGC 3664
QY 3414 cttcattaatgactggaagatctgcccctc-----cttcaactg 3452
Db 3665 ACTCCAGAAAGTCTGAAAGATGATGATCCAAAACAGTTCAGGAAGGACACAGCTCTCG 3724
QY 3453 ccttaccgaagtcacacccctcgtaaccggtctgtccagacaagccatcaagtatccg 3512
Db 3725 ATGCAACACGAACGACCGACGAAGGAAGCCCAACCCAGTCTCTCCAGCAGACACATCTCT 3784
QY 3513 gaacagagcgaagggcgagcagcttccacagcttcca-----cgaagggaaatgagga 3569
Db 3785 TCGGCGAGTCCCTGGAAAGACAGCTCCAGAGTGCAGAGTTCGAAGCGCGAGTATAGAAC 3844
QY 3570 gaagtctatggtatagagagagaatccgaatgaaacttaa-----ca 3617
Db 3845 AACCAAGCTTATGATGAGAGAGGAAAGAGACAGATGAAAGTGAAGTTCAACCGATTTA 3904
QY 3618 aatacaaacaggttagccggatccagaccacagactgagcaagatcaccatccgaagatc 3677

Db 3905 AGCAGACACAGGGTCTGGGTCTACTCAAGAACTTGACACGACGAATACTCCAAGTACA 3964
QY 3678 gctcagatggtatcccatagaggggcagatctgttccatcaatcctctgagactg 3737
Db 3965 ACATACATTAACATCAGTACAGAAAGCTGTATAACCGCTGTGCAAGTCTTCCAGATAGTG 4024
QY 3738 atgtatgatatcatctggttcaagagactagtagtctgtccttcaagacacaa 3797
Db 4025 ATGTGATGATGTGTCCGCTTCCAGAGCCAGCAGTACTTACCTTACGCTTACGACGACAA 4084
QY 3798 gctacatgtccctccaatcaagagcgccgagaggaacagaaatcagtgcttccat 3857
Db 4085 GCTTATGTCAGAGCAGAGTGTGAGCGCCCAAGG-----GTAGATCATGTTATTTACCC 4138
QY 3858 ccaaaatgcaaaacagacagatggtgcgtgtcggggaagaacttgcaccaaaagccacga 3917
Db 4139 CCAAAATGCAAGCAACAGATGCGATGCGACTTCAGAGAGAGCATCATMAAGACACAGTG 4198
QY 3918 tcaatgagacatgtctcaactgagagaalagcgcagcagctccgacatgcaatg 3977
Db 4199 TAAGTGAGAGATATATACCTTGGACGTATATACCGTATACCGATCGGACACGCGCTAG 4258
QY 3978 gcgccttgggtaccagtggtcaagaagcggtcatctagcatgtgggccaatgtgacta 4037
Db 4259 GTACGCTCGAGCGCGGTGGAAGAAAGAAAGAAATCCAGCCTGAGAGGCCAAAGTGGTAGCCA 4318
QY 4038 ttgtgtgtcttcaaggaagtcgacgtgctcctcaatcagccaaacccgaagagag 4097
Db 4319 TTG-----TGTTCCAGAGAGCAGAGAGCAGTACAGCTTACGCAACAGAGTGGGCC 4372
QY 4098 gtaaaagctacggagcactgtctcagagaagcagcgagaccggctcagctgtagatga 4157
Db 4373 ACNAGAGTTGAAGAACCCATCCAGAGAGTACGGAACAAAGAAATGGCAGCTGAATTC 4422
QY 4158 ggaactgtgaccccgccagcgccagcgaggaatccacagatggtcagatgaacagctata 4217
Db 4433 GGAA---GATGGTAGACAGAGCCGAGCGGAGTCCAGATGGCAGATCAACAGTTATA 4489
QY 4218 gctcgaagaagaatctgacttcccttgggttccgctgtgcctctgaacgcagtgtcgtg 4277
Db 4490 GCTCGAGAGAACTTGATATTTCTGAGTTGAGTAGGACCCGACAGTCAGTTCACTG 4519
QY 4278 atttcttgatggtcgtggggtcctctcagctagtgagacccagaccctgtactcctg 4337
Db 4550 ATTTCCTTGAATGGGTTGGGACACAGCAGCTCGTGGCCCTCAGACCTCCGCCACCCCG 4609
QY 4338 caatggttgcattcagtggtgaaatgatatgataaaaggaacagcttggaggttagaataca 4397
Db 4610 CCATGGCGCATATCCAAATCGGATGGATGAAGAGGGTCAAGTTGAGGTTA 4469
QY 4398 tccgggcgcggcgtctgtgtataaaccaagttccaaagacacgtccagaccgtatgtca 4457
Db 4670 TCAGAGCCCGGAGCCCTTACCAAAAACCTGTGTTCAATCTTACACCCGCTCTATGTGA 4729
QY 4458 aggtgatctgttgaacaaagaggtgtgcatagccaaaagaaacccaaggttgcgagaa 4517
Db 4730 AAGTATATCTTTTGGAAATAGAGAGCTGTATTCGCAAAAAGAAAGACAAAGATTCCAGAGA 4789
QY 4518 agacccttgagccctgtacacgaagcttcttccctcgtgagagagcccccagggaggg 4577
Db 4790 AAATCTCGATCTTGTATACAGAGTCCCTGGTTTGTATGAAGATCCACAGGATTAAG 4849
QY 4578 tgttaacagatcatctgtctggggagatatagttgtatgatatcaaaatccctttagagag 4637
Db 4850 TTCTTACGAGTATGTGTGGGAGTATGAAAGATGAAAGACCAAAATGCTTATATGGGTG 4909
QY 4638 tggccagataactcttagatgaactgaactatcaacatggtatggtatggttcaaac 4697
Db 4910 TGCGTCAAAATCTTGTGGAAGAACTTGATCTATCCAGCATGTGTATGATGATTAAT 4969
QY 4698 tcttccctcttccctcctagtagatccaactcgcacactctgacaagaagagctccc 4757
Db 4970 TGTTCCTCTCGCTCCTACCTGATGATCCACTGCTCCCTGACCCGCGGCTTCC 5029

Oy	1062	ccaagagatatacttgatgatgataaagtgtggaagcaagatgaaatgaaacaaa	1121
Db	1462	TCACACCGGCC-----TGAGAAACGGCAGTATGGCCAGAGTATGGCCGGAAGACCAAGAAAGAGA	1516
Oy	1122	ggagagagagagaatataccagcaagctacaaagtgatccaaactctgcccgtatccg	1181
Db	1517	GGAAAGGAGAGAGACTACGACTAGATACCCAGAGACCTTAACCTGGCTCGCTAACCCGG	1576
Oy	1182	taaaagccaaacccctaacgaagaanaaatgcatccatccacgtctgaagtggtccaaggaagac	1241
Db	1577	TGAAGGCCCCGCCAAGAGGACGAGATGGCATGCACGCCCGGATGTCTCCAGACGAAGC	1636
Oy	1242	atgagaagaagcacagtatgttcttttgcaaacgcgtgaactagaagaattccagattt	1301
Db	1637	ACGACCGCGCGCCACAGCCAGCTGGGGCTCCGCCAACCGAGGACACTCCGCCCGCGCGG	1696
Oy	1302	cctctctaaagatgtagatagaccatcaagcaagaatctgtatctgaacgtagaagctgcaa	1361
Db	1697	CTGAGGCCACGGCGGGCAAGCGCGCGGCCA-----CCGGCA	1735
Oy	1362	tggaaaaccaacgactcgtatctaatvgaaagaactcgcgaggtctaaggacaagtctt	1421
Db	1736	GGGTCTCTCCCGGAGTCCCGCGCGCACGGCGGGCGCCGCCAGCCTC-----	1786
Oy	1422	atccaaagaagacctcaaatatagctcctccaacccctcggcgagaccctataacgcttg	1481
Db	1787	-----CCACCGAGCACGGGCCACCGCGCGCGGCCAGCGCGCGGCCGCGAG	1834
Oy	1482	atagaccagaatbgaagcgcgctgactccctacgcgaacaacacactatagatccaggt	1541
Db	1835	AGCCACCCAGCGCGCGCTCCCGGACCCCTCTGTAAAGAGGGCCCGCTGGACCGGGCT	1894
Oy	1542	ctgc-----tgtgagaaagcaagcgagaaanaaattgaaacacgtttaagaaatgatt	1595
Db	1895	CGGCGCTTCTTGCCCAAGGCCAACGCGGAGAGAGGCGGAGACATGCTGGCGGAACGACT	1954
Oy	1596	ctttgattcaagaccgaactccgagtcagtgagagcgcccccaacgaagctcataatcca	1655
Db	1955	CGGTAGCTTCGATCAGTCCGATCCGTGGCGCATCCCGGCCAACGCTCACCGGCCCA	2014
Oy	1656	agaaaagagtaaaatgagcgccgaagtttactatgaagcgctcggagagagactgcatcca	1715
Db	2015	AGCGGGAGGCAAGAGACGTACAGTGTGCTGTAGACACTCGAGAGAGAGGGGCTGCCA	2074
Oy	1716	caactgagatacaagctgcatgattgtgagacttgaaagcgagagtgtagtgagaaag	1775
Db	2075	CACCGGAGTACGACGACTGGAGGACTGTGAGCTGTGAGACGAGAGCGTGTAGCGAGAAAG	2134
Oy	1776	ggagacagtcaaaagggaaaaaagaanaactagtgagcagggagatttgcgattctaca	1835
Db	2135	GTGACTT-----	2141
Oy	1836	ccaggtctgagagacaaaagaagaagtactatgtgtgcacctcttvgaaagagatt	1895
Db	2142	-----GGATT	2146
Oy	1896	tggaatggtctgagcctcaagataaagaactctggtgtagataactgttagtagcaaaccc	1955
Db	2147	ACTACTGTGTGATGCCGCCGCTGTATTCTTACCAAAAGAACAAAC-----CATGCCCA	2284
Oy	1956	ttaacgagagcatagcctatgtagtaagacacctctgaccttgccagccatccaagaatg	2015
Db	2199	-----TTGCACTCCTGTACGCTGGCAGCGCTTAAAGAGG	2233
Oy	2016	gagatcgactaatgtgtcgtatcttattatnaatbaagcgctttaaagaatggaagtgtacct	2075
Db	2234	GAGATCGACTAATCGCGCGTGTATTCTTACCAAAAGAACAAAC-----CATGCCCA	2284
Oy	2076	gagactcagagcaatgtgttgacttaaaagttgttagagagaaagaatgaactgaatcagctc	2135
Db	2285	AAGAAATCAGGTCTATTATTGGCTCTGAAGGTGTGTTGGAGGAAATATGACGACTTAAGGC	2344

OY	2136	gacfttgycaatttatlaccaaaglaaaaaaggaagtttagctgtatctgttgaagcattc	2195
Db	2345	gcccttggcgttttcataccaaataaagaaggccacccttcgcagacgtctgcgacacc	2404
OY	2136	ttagaacggatgagaagttcttgaatggaaatggagagctatttgcagaaggacacatttg	2255
Db	2405	taaaGCCAGGGGAGCAAAATCCTTAAGTGGAAATGGTAAACCCCTGCCGGAGCAAAACG	2464
OY	2256	aggaagtttacacaattctctagaatccaaaccttgacaaccaaatgttagctgttgttt	2315
Db	2465	AAGAAGTTTACAACTATTTCTTAGATATCAAAATATGAAACCTCAAGTTGAGATTTATGTTT	2524
OY	2316	caaggccaattggagatattcctgaataacttgaagaagcgtacgcagtcgacaacttgatcca	2375
Db	2525	CAAGGCCATATTTGGTACATCCCAAGATCCCTGAGAGTTCCCATCTCCCTCGAGTCA	2584
OY	2376	gtctctagctcatltaactcaaaaaatggaacgylccctctataltccgttaactcaccga	2435
Db	2585	GTTCAAAGTTCCCTTGATATCTCGAAAAATGGAAGGCCCTTCTATTTCTGTATTTTCTCGAA	2644
OY	2436	taagtcctggcactgtcgtgaggaatgcccgcagttcttctcggacaagtttcaataaac	2495
Db	2645	CCAGCCCTTGACCTCGAAAGATGGCCCAACAGTTTACACAGGGCAACTCTCAGTGAAGC	2704
OY	2496	tatgtgttgaacaagtttggccaacgttgatatgtatcaatttggagcaaaagatctcc	2555
Db	2705	TATGTTATATTAAGTGGGGCACACAGCTGATTTGTAATGTTCTACAAACACAGATCTAC	2764
OY	2556	cttccaggaagaatggagagccaagaatccttaagttaagaattacttctctccaaata	2615
Db	2765	CCCCCTGAGTAGATGGCCCTCCAGGAATCCCTATGTAAAAATGTATTTTCTTCAGATA	2824
OY	2616	gaagtgtataaaaataagagaagaacaacaacagtcacaagaaaacttltggaaccacaatgga	2675
Db	2825	GAAGCGACAAAATTAAAGGAGACCAAAACAGTAAAGAACTTCTAGAGCCCAAAATGGA	2884
OY	2676	acagaacttcatltaattctcctgtccacgaagagaattccgtgaacgaatgctggaaa	2735
Db	2885	ACCAGACATTTGCTACTCTACACAGTACATCTGTAGAATTTTGGAGGCGAAATGTTAGAGA	2944
OY	2736	ttaacctttggagatcaagcttagaggttctcggaagaagaagacgaattcttagagagaatt	2795
Db	2945	TTTACCGTGTGGGACCAACCAGAGTACAGACGAAGAAGATGTAATTTCTTGAGAGATCC	3004
OY	2796	taactgaattggaaacacagcttctctagatgatagagccgcacgtgtataagctctgcagacc	2855
Db	3005	TCATAGAGTTTGGAAACAGCCCTTTTATGATGATGAGCCCATTTGGTATTAACCTCCACATAC	3064
OY	2856	atgatgtctcctaatcttgccaatcccttcgccttccccaatctcgcgccggagacgaatcc	2915
Db	3065	ATGACGAATCTTCTACTACTCTCTGCTTAGCCATTCACCGTTCAATGCCACAGGCCATATATTC	3124
OY	2916	atggaagaagcccaacgcgcgacgtctgcaaaagtcogaagaataagtgcagtgaagtgt	2975
Db	3125	ATGGAGAGAGGTCCAGCAAAAAGCTATCAAAAGATCTCAGCGAAATCAGTAAAGTGAATCTC	3184
OY	2976	ctgaactacgactcgcgaggaatggcgttggagtagtgtc-----agattatcgacacaatg	3029
Db	3185	CAGATTATGAGGTTGATGATGATGTATTTGAGATGAGTCCCTCCAGTGGGTTTAAAGACTAGTG	3244
OY	3030	gcccgcgacttccaagctccacagcttgcgtgtgcgcagacaagaatcatgtgtcatcaatcatc	3089
Db	3245	CTAGAGAGAGTAAAGCAACACACAGTAAACAGTGGCCAGCAAAACAAGAACTACATCACC	3304
OY	3090	gctcaaccaatcaggtctctccatcatcgagttagatgttataaggaagaacaaggttcagtgcgc	3149
Db	3305	GCTCAGCTTCGTTGCTCTCTCTATCGCCGCGATGATCAAGGAAAGCCCTGCTTACAGTTTAC	3364
OY	3150	ctaaatgcccctctcctcaaaaggaatgtcgaa-----caagggcacccgagaggaacgtgtcta	3206
Db	3365	CAAAATGTGCATTACAGAGAGACCTTAAATGAATTTCAATCAACCAAGAAAGTCAAGTCTCTC	3424
OY	3207	ctggccattatacaacaatttagccgaa-----tggatagaacccgtgtcatgtatg	3267

Db	3425	CAACCCGACACCATGATGCCCTCCCGAAGCCCGGCCATCAGATCCAGCATATGTGAAA						3484
Qy	3258	accactactcttcacatagacagagga						3285
Db	3485	GTCATATTTCGTGACAGCCAGACAGTGGACTTCTCATGTGCCACAGCAAAACGAGAC						3544
Qy	3286	-----						3285
Db	3545	GAATGTGAGAAAACCTACACATGACCAGTGAACCTGCAGCCCTCTTGACAGGGCTAGGA						3604
Qy	3286	-----						3300
Db	3605	GTGTAGTACCAACTGCTTGAGACCAAGATACTAGTTTGCACTTTCACAGAACGAGAAAAGC						3664
Qy	3301	-----						3300
Db	3665	ACTCCAGAAAGCTGAANAATGTAGATCCAAAACAGTCTAGGAAAAGCAGACGCTCTG						3724
Qy	3301	-----						3329
Db	3725	ATGCAAGACAGGGTTCTCCACCATGCTTCTTGAGAGGGGATACGCAACCCCAAGACAA						3784
Qy	3330	cagaacacacgcctctcccaagvcygaaccaaccaaccgctccgaatctctcgaacgtctg						3389
Db	3785	CCGATCAACCGGTCGTTAGGGGAAATATCCCATCTGTTACCGGTGAGAGGACTCTA						3844
Qy	3380	atacaaacctctagagtcgaatgcattcaatagctgaagaaatctgcccctctc---						3446
Db	3845	GTGTAGAAACCTGTGTCTTATGTACACCACTTCCCGGAGGGGTGGGCCCACTTCTC						3904
Qy	3447	caactgccttacgcaggtctcaacctcgtaaccggtctgtctcagacaagccatcaagta						3506
Db	3905	CACCTTCTTGTGACAAAGAACGCAACCGCAAGAAAGCCCAACCCAGTCTCTTCAGCAGACA						3964
Qy	3507	ctccgggaacaggaegaagggcccgacagcttccacagctcca---ccaaggygaacat						3553
Db	3965	CATCCTTCGGCAGTGCCTGTGGAGAGCAGCTCCACAGGTGCCAGTTGCAAGGGCGAGTA						4024
Qy	3564	tggagagaagtgtctatgatatagagagagaaatcgccaaatgaaacttaa-----						3615
Db	4025	TAGAACMACCAAGCTTAGTAGTGTGGAAGAGCGAAGCAGACAGATGAAAGTAAAGTTACC						4084
Qy	3616	---caataacaaacagagttagcccggaatcagaaccacagactcgaggaagattacatctga						3671
Db	4085	GATTTAAGCAGACACACAGGGGTCTGGGTATGATGAAGAACTTGACACAGACAAATATCCCA						4144
Qy	3672	agatcgcctcagatctggatccacatagaggggcagatactgttccaactaaatcctcgg						3731
Db	4145	ACTAACAACTATCAATAAAGATCACTACAGAAAGCTGTATTAACGGGTCTGCCAAGCTTCAG						4204
Qy	3732	acaagatgabtgaagiatgfatctcgsggtttcaagsgctagtagtgccttcgtttcgga						3791
Db	4205	ATAGTGTATCTAGTGTATGTGTCCGCAATTTCAGACCCACACAGTACTTCACGGCTTCAGCA						4264
Qy	3792	gcacaagctacatgtctcgctccaatcagagcgcgccgagaggaagaaacaggaanaatcagtgct						3851
Db	4265	GCACAAAGCTTATGTGTAGAGCAACTGTGAGCGGCCCAAGG-----GTAGGATCAAGTTCAT						4318
Qy	3852	ttacaatccaanaatgcaaaacagacagatggcggtgtctcggggaaagaaacttgacaaaagca						3911
Db	4319	TTAACCCCCCAAAATGCAAGCAGACGAGTGGGCGACTTCAGAGAAAGCCATCATCAAGAGCA						4378
Qy	3912	ccagatcaagtgagagacatgctctcaatctgagaaagaaatgacgycgaagcagctgcacactg						3971
Db	4379	CCAGGTGAAGTGGAGATATATATACCTGGAACGTAAATGACGGTAGCGACGTGGACACGG						4438
Qy	3972	cagttgggccccctgggtacacagtggcaagaagcggcgatctagcatctggggccaaaatg						4031
Db	4433	CCGTAGGTACCGTCCGAGCCGGGTGGAAAAGAAACGAAGATCCAGCTTAGCGGCCAAATGG						4488
Qy	4032	tagctatgttggtctctcaagaaaagtcgcagtgcttcaactcaagcnaaacggaag						4091

Db	4439	TACCATTTG-----TGTGTGAAGAAGACGAGACGCTACACGCTACGACGACGACGAGT	4552
OY	4092	gaggsgtaaaaaagctacgagcaactgltcagagaagcaagagaccggctagcaagt	4151
Db	4553	CGGGCCACAGAAAGTTGAAAAACACCATCCACGAGGAGTACGGAAACAGGAATGGCAGCTG	4612
OY	4152	agatagagacgtgattgaccgccgagccgagccggaaatccaaagatggcagatgaataa	4211
Db	4613	AAATCCGGGAA--GATGTGAGACACAGCCGAGCCGGGAGTCCACGAGTGGCATCATCAAC	4669
OY	4212	gctatagctcggaaagaaatctgacitcccttcggggctccgctcggctctgacagccagt	4271
Db	4670	GTTATATAGCTCGGAGAAACCTTGATATTTCTCTGGAGTTGGAGTAGGACCCGACAGTCAGT	4729
OY	4272	tcagtgattccctgagatggctctgggcccctgctcaagctagtgagacgccaagaccctgcta	4331
Db	4730	TGAGTATTTCTTGATGTGGGTTGGGACACAGCGACGCTCGTTGGCCGTACAGACCTCGCCA	4789
OY	4332	ctctctgcaattggtgacatctcaagtgaggaaatgataaataaaaggacagctggaagttag	4391
Db	4790	CCCCGGCCATGGGCGATATCCAAATTCGGATGGAGGATAAAGAGGGTCAAGTTGGAGCTTG	4849
OY	4392	aaatcatccgggcgcgcgcgcctctggttaaaacaggttccaaagacacgctgcagccagct	4451
Db	4850	AGCTTATATCAGACCCCGAGCCTTTACCAAAAACCTGTGTTCCAAATATTACACCCTGCTCCT	4909
OY	4452	atctcaagaatgatatctgtctagacaacagagatctgcaatgaccaaataaaagaaaccaaagt	4511
Db	4910	ATGTGAAGATATATCTTTTGGAAATATGGAGCGCTGATTTGGCAAAAAAGAACGAATTG	4969
OY	4512	cgaagaagaacctggaagccctgtacacagcagctcttgccttcgagagagaccgccag	4571
Db	4970	CACGGAAAACCTCGATCTCTTTGATACACAGTCCCTGTGTTTTGATGAAAGTCCACAGG	5029
OY	4572	ggaagggtgtaagatcatctgtctcggggagatatagtlogtatagtatccaaatcttta	4631
Db	5030	GTAAGATCTTCAGGTGATGTCTGGGGTACATGAGAAATGAGCCACAAATATGCTTTA	5089
OY	4632	tggaggatggccagatattccttagatggaactggaacttccaaatcagtgatctgagtggt	4691
Db	5090	TGGGTGTGGCTCAAACTCTGTGGAGAACTTGATATCCACATAGTGATTTGGATGTG	5149
OY	4692	tcaaatcttccctccctcttccctcctagtagatccaaactcggacacttgaacaagaag	4751
Db	5150	ATAAATTTGTCCTCCGCTCCTCACTGCTGGATCCCACTCCGCTCCCTGACCCGCGGG	5209
OY	4752	cttcccaatctgctctcgaaagttctaacggacctcttactctgtgttatagcaa	4807
Db	5210	CTTCCCAATCATCTCTGGAAAGTCTGTCGGGGCTCTCTGTGCATCCGTCATATGATTA	5265

Search completed: November 21, 2001, 18:56:01
Job time: 10687 sec

Db	3425	CAACCCGACACCATGATGCCCTCCCGAAGCCCGGCCATCAGATCCAGCATATGTGAAA						3484
Qy	3258	accactactcttcacatagacagagga						3285
Db	3485	GTCATATTTCGTGACAGCCAGACAGTGGACCTTCTATGTGCGCCAGACGAAACGAGAGAC						3544
Qy	3286	-----						3285
Db	3545	GAATGTGAGAAAACCTACACATGACCAGTGAACCTGCAGCCCTCTTGACAGGGCTAGGA						3604
Qy	3286	-----						3300
Db	3605	GTGTAGTACCAACTGCTTGAGACCAAGATACTAGTTTGCACTTTCACAGAACGAGAAAAGGC						3664
Qy	3301	-----						3300
Db	3665	ACTCCAGAAAGCTGAANAATGTAGATCCAAAACAGTCTAGGAAAAGCAGACGCCCTCTG						3724
Qy	3301	-----						3329
Db	3725	ATGCAAGACAGGGTTCTCCACCATGCTTCTTGAGAGGGGATACGCAACCCCAAGACAA						3784
Qy	3330	cagaacacacgcctctcccaagvcygaaccaaccaaccgctccgaatctctcgaacgtctcy						3389
Db	3785	CCGATCAACCGGTCGTTAGGGGAAATATCCCATCTGTTACCGGTGCAGCGAGACTCTA						3844
Qy	3380	atacaaacctctagagtcgaatgcattcaatgaatgcgaatgcgaatgcgaatgcgaatgcga						3446
Db	3845	GTGTAGAAACCTGTGTCTTATGTACACCACTTCCCGGAGGGGTGGGCCCACTTCTC						3904
Qy	3447	caactgccttacgcaggtctcaacctcgtaaccggtctgtctcagacaagccatcaagta						3506
Db	3905	CACCTTCTTGTGACAAAGAACGCAACCGCAAGAAAGCCCAACCCAGTCTCTTCAGCAGACA						3964
Qy	3507	ctccgggaacaggaegaagggcccgacagcttccacagctcca						3553
Db	3965	CATCCTTCGGCAGTGCCTGTGGAGAGACAGCTCCACAGGTGCAGTTGCAAGCGGCGATA						4024
Qy	3564	tggagagaagtgtctatgatatagagagagaaatcgccaaatgaaacttaa						3615
Db	4025	TAGAACMACCAAGCTTAGTAGTGTGGAAGAGCGAAGCAGACAGATGAAAGTAAAGTTACC						4084
Qy	3616	-----						3671
Db	4085	GATTTAAGCAGACACACAGGGGTCTGGGTAGTCAAGAACTTGACACAGACCAATATCCCA						4144
Qy	3672	agatcgcctcagatctggatccacatagagggcagatagcttccaaactaatctcgg						3711
Db	4145	ACTAACAACTATCAATAAAGATCACTACAGAAAGCTGTATTAACGGGTCTGCCAAGCTTCAG						4204
Qy	3732	acaagtgatgaagtagatgtatctcgsggtttaaagsgctagtagtgccttcgtttcgca						3791
Db	4205	ATAGTGTATCTAGTGTATGTGTCCGCCATTTCAGACCCACACAGTACTTCACGCCCTCAGCA						4264
Qy	3792	gcacaagctacatgtctcgtcccaatccagagcgcgccgagaggaagaaacaggaanaatcagtgct						3851
Db	4265	GCACAAAGCTTATGTGTAGAGCAACTGTGAGCGGCCCAAGG-----						4318
Qy	3852	ttacaatccaaaatgcaaaacagacagatgvgcgltctcggggaagaaacttgacaaaagca						3911
Db	4319	TTAACCCCCCAAAATGCAAGCAGACGAGTGGGCGACTTCAGAGAAAGCCATCATCAAGAGCA						4378
Qy	3912	ccagatcaagtgagagacatgctcactctgtagaagaatagtcgycgaagcagctcgcacactg						3971
Db	4379	CCAGGTGAAGTGGAGATATATATACCTGGAACGTATAGCGGTAGCGACGTGGACACGG						4438
Qy	3972	cagttggcgccctctgggtacacgttgcgaagaagcggcgatctagcatctggggccaaaatg						4031
Db	4433	CCGTAGGTACCGTCCGAGCCCGGTGGGAAAGAAACGAAGATCCAGCTTAGCGGCCAAATGG						4488
Qy	4032	tagctatgttgctctctcaagaaaagtcgcagtgcttccactaacgcaaacggaag						4091


```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
```

```
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Db 600 AGTCGTACACATAGTGCCTAGCAAGG 627

RESULT 2
AA476826 590 bp mRNA EST 08-AUG-1997
LOCUS AA476826
DEFINITION IMAG:784752.5', similar to WP:TI10A3.1 CE04931 C2 DOMAIN OF PROTEIN
KINASE C ;, mRNA sequence.
ACCESSION AA476826.1 GI:2205037
VERSION AA476826
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 590)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gaisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wyllie,
T., Waterston, R. and Wilson, R.
Washu-Merck Est Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 495.
Location/Qualifiers
1..590
/organism="Homo sapiens"
/db_xref="GDB:5982370"
/db_xref="taxon:9606"
/clone="IMAGE:784752"
/clone_1lb="Soares_Total_fetus_NB2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCTTAATTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 188 a 135 c 129 g 138 t

ORIGIN

Query Match 9.0%: Score 446.4; DB 8; Length 590;
Best Local Similarity 89.4%: Fred. No. 1.6e-114;
Matches 504; Conservative 0; Mismatches 56; Indels 4; Gaps 2;

Db 184 AGGAGTCTCATAGCCAAAAGAAACAAAGTGGCAGAGAAAACGTGGACCCCTT 243

QY 4536 accagagctcttgccttcctgagagagagccccagggaggggtgtacagatcattgtc 4595
|||||
Db 244 ACCAGAGCTATTATCTTTTGGAGAGAGTCCACAGGAAAAGTTTACAGATCGTCC 303
QY 4596 gggagagattagtcgtatgatgatacaaatcccttatggaggtggccagatcttag 4655
|||||
Db 304 GGGAGATTATGGCCGATGATCACAATCTTTTAAGAGGTGGCCAGATCTTTAG 363
QY 4656 atgaactggaactatccaacatggtgatlgatggttcaaacctctccctctcc 4715
|||||
Db 364 ATCACTAGAGCTATCCAAATATGATGATCGATGATGATCAACTTTTCCACCTTCCTCC 423
QY 4716 tagtagatccaactcggacactcttgacaagaagagcttcccaatcgtctctggaagt 4775
|||||
Db 424 TAGTAGATCCACACCTTGCGCTCTGACAGAAAGAGCTTCCCAATCATCTTGAGAAATT 483
QY 4776 ctaccgagctcttactctcgttcacatagaaact---ataaactggtgtcaacaacc 4832
|||||
Db 484 CAACTGACCTTCTTACTCTTCGTTACATGACGCTGTATAAAAAATTGTTGTACACGACAC 543
QY 4833 agcgatacaaaaacccaagaagaaa 4856
|||||
Db 544 AGCGTTACAAAAMAAAAAMAAAAA 567

RESULT 3
AA602733/c 523 bp mRNA EST 08-OCT-1997
LOCUS AA602733
DEFINITION n17604.s1 NCI-CGAP_B-3 Homo sapiens cDNA clone IMAGE:1116630 3'
similar to WP:TI10A3.1 CE04931 C2 DOMAIN OF PROTEIN KINASE C ;, mRNA
sequence.
ACCESSION AA602733 GI:2436667
VERSION AA602733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 523)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: c9apbs-f@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbrrp/image/image.html
Insert length: 1223 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 449.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1116630"
/clone_1lb="NCI-CGAP_B-3"
/sex="female"
/tissue_type="breast tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: breast; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Ductal breast tumor. 5' adaptor sequence: 5'
GAATTCGCCAGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 0.9 kb."

FEATURES
source

Db	294	TTCCCTTGATGATTTGGGGCCACCCAGCTTGTTCGGCCGCCAAACCTCGCCACCCACGCC	353
Qy	4340	atgggtacatcaccagtgctggaaatgtagtaaaagaaggaacagctggaggtagaaatc	4399
Db	354	ATGGGCATATTCCAAAATGGGATGGAGATAGAAGGCGTCAGTTGGAGGTTGAGTTATC	413
Qy	4400	cggagcgcgcgcccttggtgtaaaccaggtgccaaagacatgcgcagaccogtatgtcaag	4459
Db	414	AGAGCCCGAGACCTTACACAAAACCTGGCTCCAAATCTACACTCTCTCCCTGTGTGAA	473
Qy	4460	gtgtatctgttagaacaagggatctgcataagccaaataaaacccaagttggagaaag	4519
Db	474	GTTATCTTTTGGAAAATGAGACCTGTATGTCAAAABAAGCAAAATGACAGGAA	533
Qy	4520	acccttgagcccttaccagcagctctgtcttcgcagagagagcccccagggagagt	4579
Db	534	ACTCTCGATCCCTTGTATACGACAGTCCCGTGTGTTTGTGATGAAAGTCCACAGGGTAAAGTT	593
Qy	4580	ttacagatcatctgtcgggagattatgctgatalgatacacaactccttlabygagtg	4639
Db	594	CTTCAGCGATTTGCTGGGGCTACATATGAAGAATGACCAAAATGCTTTATGGGHTG	653
Qy	4640	gcccaagatactctagatgaactggaactatccacaatggtgattgtagatgtccaact	4699
Db	654	GCTCAAACTTGTTGGAGAACTTGATCTGTCACAGCATGCTGATGTGATGAATTAATG	713
Qy	4700	ttccctctctctccctagtagatc	4725
Db	714	TTCCCTTCATCTCCCTGGTGAATCC	739

RESULT	12
LOCUS	BE864173
DEFINITION	BE864173 359 bp mRNA EST 29-SEP-2000
ACCESSION	U1-M-BH1-amy-h-08-0-01.r1 NIH_BMAP_M_S2 Mus musculus cDNA clone
VERSION	U1-M-BH1-amy-h-08-0-01.5', mRNA sequence.
KEYWORDS	BE864173 BE864173.1 GI:10384953
SOURCE	EST.
ORGANISM	house mouse.
REFERENCE	Mus musculus.
AUTHORS	Enxaiyola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	1 (bases 1 to 359)
JOURNAL	Bonaldi, M.F., Lennon, G. and Soares, M.B.
MEDLINE	Normalization and subtraction: two approaches to facilitate gene discovery
COMMENT	Genome Res. 6 (9), 791-806 (1996)
	9704447
	Contact: Chin, H
	National Institute of Mental Health
	6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
	20892-9643, USA
	Tel: 301 443 1706
	Fax: 301 443 9890
	Email: mestr@mail.nih.gov
	cDNA library Preparation: M.B. Soares lab Clone distribution:
	Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
	should be noted that Bento Soares is generating a small number of
	additional specialized non-redundant arrays of BMAP cDNAs whose
	availability will be considered under appropriate and limited
	collaborative arrangements
	Seq primer: M13 Reverse.
FEATURES	Location/Qualifiers
SOURCE	1..359

```

FEATURES
source
location/Qualifiers
1..359
/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone="U1-M-BH1-amy-h-08-0-01"
/clone_11b="NH-BMAF-M1S2"
/dev_stage="27-32 days"
/lab_host="DH10b (life technologies)"

```

```

/note=Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_M_S2 library is a subtracted library derived from
NIH_BMAP_M_S1, which in turn is a subtracted library
derived from a mixture of normalized libraries from ten
regions of the mouse brain (cerebellum, brain stems,
olfactory bulbs, hypothalamus, cortex, amygdala, basal
ganglia, pineal gland, striatum, hippocampus). The driver
used for subtraction consisted of a pool of 5,000 clones
from the NIH_BMAP_M_S1 library and a pool of 2,000 clones
obtained from non-normalized and normalized mouse brain
spinal cord libraries."

```

Query Match	7.0%;	Score 347;	DB 141;	Length 359;
Best Local Similarity	99.7%;	Pred. No. 1.6e-86;		
Matches 358;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;

QY 4167 tgaccgcgcaagccagcgcgaatccacaga tggcgcatgaa cagctata gctcggaa g 4226
|||||
Db 1 TGACCCGCCAGGCACCGGGGAATCCACAGATGCGACATGAACAGCTAATAGCTCGGAAG 60

Qy	4227	gaatctgatctctccctcggggtgcgcgcctgctcctgacagccagttcagtgattccctgg	4286
Db	61	GAATCTGATCTTCCCTGGGGGTCCGCTGCGCTCTGACAGCCAGTTCAGTGTTCCTGG	120

QY 4287 atgagctgggcacctgtcagactagtgtggacgccagaccctgctactccgtcaatgggtg 4340
 |||||
Dd 121 ATGCGCTGGCCCTGCCTCAGCTATGTGGACGCCAGACCCTGGCTACTCCGCAATGGGTG 180

0y 4347 acattcagtggtggaatgatgtgataaaagaagacgcggaggttagaatcatccggcgcc 4408
|||||
|||
Db 181 ACATTGAGGTGGCATGATGCATAAAAGGACACGCTGAGGTGAATTCATCCGGCGC 240

QY	4407	gcggccttctggttaaacaaggttccaaga-cacgtgcagcaccgtatcgttcaagtgat	4408
Db	241	gcggccttctggttaaacaaggttccaagcactgcacgacacgtatctcaagctgtat	300

Db 301 CTGTTAGACAACGGAGTCTGCATAGCCMAAAAGAAMACCAGGTGGCGAGCAAGACCCT 359

RESULT	13
BF510067/c	
LOCUS	BF510067
	428 bp
	mrna
	EST
	06-DEC-2000

IMAGE:3086882.3', mRNA sequence.
BF510067
BF510067.1 GI:11593365

SOURCE ORGANISM	
human.	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom	
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 428)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
Tumor Gene Index

JOURNAL COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
The sequence contained an oligo-dT track that was present in the

Standard cDNA and therefore this may represent a cDNA library tail. cDNA Library Preparation: M.B. Soares Lab Clone Distribution NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=yes.

FEATURES
source

Location/Qualifiers
1. .428
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3086882"
/clone_lib="NCI_CGAP_Sub8"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; NCI_CGAP_Sub8
is a subtracted library derived from NCI_CGAP_Sub5. The
NCI_CGAP_Sub8 library had 2.5 million recombinants. A
single-stranded DNA preparation of NCI_CGAP_Sub5 was used
as a tracer in a subtractive hybridization with a driver
comprising a pool of clones from NCI_CGAP_Sub5 (IMAGE
clone ids 2732833-2737415, 3068040-3069191; 25% of the
driver population), a pool of clones from NCI_CGAP_Sub4
(IMAGE clone ids 2723592-2729326; 25% of the driver
population), NCI_CGAP_Sub6 (pool AIF-AU, IMAGE ids
2728969-2733190; 25% of the driver population), and
NCI_CGAP_Sub7 (IMAGE ids 3069192-3072238, 3081664-3084550
; 25% of the driver population). Subtraction was
performed as previously described [Bonaldi, Lennon &
Soares (1996): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.
TAG_LIB=NCI_CGAP_Lus
TAG_TISSUE=Lung
TAG_SEQ=CAAC"

BASE COUNT 105 a 82 c 97 g 144 t

ORIGIN

Query Match 6.8%; Score 336.8; DB 149; Length 428;
Best Local Similarity 88.3%; Pred. No. 1.3e-83;
Matches 378; Conservative 0; Mismatches 47; Indels 3; Gaps 1;

QY 4431 ccaagacactccgacgacgtatgcaaggtgcatctgttacaacaggagctgcatag 4490
|||||
Db 428 CCAAGACCTCGCACCGTATGTAAGTGTATCTATCTATGATTAACGAGCTCGATAG 369
|||||
QY 4491 ccaaaagaacaaaggtgaggaagaagccctgagccctgtacacagcagctctgt 4550
|||||
Db 368 CCAAAAACAAACAAAGTGGCAAGAAAACCGCTGGAACCCCTTACACAGACTATAT 309
|||||
QY 4551 ccttcgaggaagaccccaaggaggaggtgtacacatcatctgtcgggagatagtc 4610
|||||
Db 308 CTTTCGAGAGAGTCCACAGGAAAGTTTACAGATCATCTCTGGGAGATTTATGGCC 249
|||||
QY 4611 gtatggtacacaatccttattgaggtgagccagatattctttagaactggaact 4670
|||||
Db 248 GCATGATCACAAATCTTTATGGAGTGGCCACATATCTTTAGATGAACTAGACTAT 189
|||||
QY 4671 ccaacatggtatgattgattgattcctcctcctcctcctcctcctcctcctcct 4730
|||||
Db 188 CCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 129
|||||
QY 4731 cggacacctctgacaagaagagcttcccaatgctctctggaaggtcttaccggaact 4790
|||||
Db 128 TGGCTCTCTGACAGAGAGCTTCCATCATCTCTGGAAGTTTCACTTGACCTCTTT 69
|||||
QY 4791 actctgctcatagcaact---ataaacgtgtgcacacacacacgagatacaaac 4847
|||||
Db 68 ACTCTCTCTGATAGAGCTGTAAATAATGTGTTCACAGCAACCGCTTACAAAAA 9
|||||
QY 4848 agaagaa 4855
|||||
Db 8 AAAAAAA 1

RESULT 14
BE980779 456 bp mRNA EST 05-OCT-2000
LOCUS BE980779
DEFINITION UI-M-BG2-bcn-a-01-0-UI.s1 NIH_BMAP_MSC_s1 Mus musculus cDNA clone

ACCESSION BE980779
VERSION BE980779.1 GI:10649193
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 456)
AUTHORS Bonaldi,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov
Oligo-dt track not found. Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENERICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA=No.

FEATURES
source

Location/Qualifiers
1. .456
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BG2-bcn-a-01-0-UI"
/clone_lib="NIH_BMAP_MSC_s1"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP_MSC_s1 library is a subtracted library derived
from which this clone was derived, please visit our web
site at brainest.eng.uiowa.edu.
TAG_SEQ=None found"

BASE COUNT 129 a 116 c 97 g 114 t

ORIGIN

Query Match 6.7%; Score 332.4; DB 142; Length 456;
Best Local Similarity 98.2%; Pred. No. 2.3e-82;
Matches 336; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4639 ggcccaatactctttagatgacagtggaactatccaactggtgattgattgattcaact 4698
|||||
Db 2 GGCGCGTACTCTTATGATGAGTGAAGTCAATCAATGATGATGATGATGATGATGAT 61
|||||
QY 4699 ctccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 4758
|||||
Db 62 CTTCCTCCATTTCTCCCTAGTATGATCCAACTTGACACTCTGACAAAGAGCTTCCCA 121
|||||
QY 4759 atcgctctctgaaagtcctaccggaacttctaactctgcttcataagcaactataaactg 4818
|||||
Db 122 ATCGTCTCTGGAAGTTCTACCGGACCTTCTACTCTGTTCAATGCAACTATAAAC 181
|||||
QY 4819 ttgtcacacaacacgagatatacaaaacagaagaacacgagagtggaagcccttggt 4878
|||||
Db 182 TTGTCAACAACACGAGATCAAAACAGAGAAAGCAGAGTGAAGCCCTGTGT 241
|||||
QY 4879 aacactgcatgattgattgtctacagaccacgcttaggatatccaagcagctct 4938
|||||

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2001, 20:56:50 ; Search time 143.6 Seconds
(without alignments)
7854.181 Million cell updates/sec

Title: US-09-617-099B-2
Perfect score: 4980
Sequence: 1 gcttcctaggtgtgttcg.....acattgtgcctagaagaag 4980

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents,NA:*
1: /cgnl_7/prodata/1/lna/5A_COMB.seq:*
2: /cgnl_7/prodata/1/lna/5B_COMB.seq:*
3: /cgnl_7/prodata/1/lna/5A_COMB.seq:*
4: /cgnl_7/prodata/1/lna/6B_COMB.seq:*
5: /cgnl_7/prodata/1/lna/PT05_COMB.seq:*
6: /cgnl_7/prodata/1/lna/Dackfilesl.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	67.2	1.3	7218	1 US-08-232-463-14	Sequence 14, Appl
2	47.6	1.0	2244	3 US-09-094-714A-48	Sequence 48, Appl
3	47.6	1.0	2245	4 US-09-225-749-24	Sequence 24, Appl
4	46	0.9	533	6 5482709-5	Patent No. 5482709
5	45	0.9	3489	2 US-08-728-323A-1	Sequence 1, Appl
6	45	0.9	32207	2 US-08-770-379-20	Sequence 20, Appl
7	45	0.9	32207	4 US-08-757-669A-20	Sequence 20, Appl
8	43.4	0.9	397	3 US-09-253-691-3	Sequence 3, Appl
9	42	0.8	203	4 US-09-043-303-7	Sequence 7, Appl
C 10	41.4	0.8	51259	3 US-08-781-891-209	Sequence 209, App
11	40.8	0.8	477	4 US-09-135-994-1	Sequence 1, Appl
12	40.8	0.8	543	6 5273901-6	Patent No. 5273901
13	40.8	0.8	3891	4 US-09-036-315-1	Sequence 1, Appl
C 14	40.6	0.8	16442	3 US-08-781-891-208	Sequence 208, App
15	39.6	0.8	154	1 US-08-469-802B-6	Sequence 6, Appl
16	39.6	0.8	154	2 US-08-267-803B-6	Sequence 6, Appl
17	39.6	0.8	165	4 US-09-043-303-17	Sequence 17, Appl
18	39.6	0.8	168	1 US-08-469-802B-4	Sequence 4, Appl
19	39.6	0.8	168	2 US-08-267-803B-4	Sequence 4, Appl
20	39.6	0.8	171	1 US-08-469-802B-5	Sequence 5, Appl
21	39.6	0.8	171	2 US-08-267-803B-5	Sequence 5, Appl
22	39.6	0.8	195	1 US-08-469-802B-2	Sequence 2, Appl
23	39.6	0.8	195	2 US-08-267-803B-2	Sequence 2, Appl
24	39.6	0.8	234	1 US-08-469-802B-3	Sequence 3, Appl
25	39.6	0.8	234	2 US-08-267-803B-3	Sequence 3, Appl
26	39.6	0.8	3376	1 US-08-320-559-29	Sequence 29, Appl
27	39.6	0.8	3376	3 US-08-545-860D-29	Sequence 29, Appl

28	39.6	0.8	3376	5 PCT-US94-04496-29	Sequence 29, Appl
29	38.4	0.8	2301	1 US-08-306-691B-23	Sequence 23, Appl
30	38.4	0.8	2301	5 PCT-US93-06251-78	Sequence 78, Appl
31	37.8	0.8	289	4 US-09-007-005-17	Sequence 17, Appl
32	37.8	0.8	289	4 US-09-244-796-17	Sequence 17, Appl
C 33	37.8	0.8	1441	1 US-08-136-277-18	Sequence 18, Appl
C 34	37.8	0.8	1441	2 US-08-479-403-18	Sequence 18, Appl
C 35	37.8	0.8	1441	3 US-08-835-734-18	Sequence 18, Appl
36	37.4	0.8	1994	2 US-08-933-750C-58	Sequence 58, Appl
37	37.4	0.8	1994	3 US-09-234-613-58	Sequence 58, Appl
C 38	37	0.7	4411529	4 US-09-103-840A-1	Sequence 1, Appl
39	36.6	0.7	921	2 US-08-795-475-2	Sequence 2, Appl
40	36.6	0.7	1308	2 US-08-795-475-4	Sequence 4, Appl
41	36.2	0.7	5285	2 US-08-609-049A-29	Sequence 29, Appl
42	36.2	0.7	5285	4 US-09-170-996-29	Sequence 29, Appl
43	36	0.7	3337	1 US-08-072-610-1	Sequence 1, Appl
44	36	0.7	3337	2 US-08-719-822B-1	Sequence 1, Appl
45	36	0.7	3337	4 US-09-092-458-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESS: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: PTZ9pt-F15
: US-08-232-463-14

Query Match 1.3%; Score 67.2; DB 1; Length 7218;
Best Local Similarity 6.4%; Pred. No. 1.1e-09;
Matches 27; Conservative 232; Mismatches 165; Indels 0; Gaps 0;

1648 taatccaagaagaggtataatgctccaggttctcactgagcaactggagagagagct 1707
1437 TACRR 1378
1708 ggcctccacactgatacactgctgagtgatgagctggaagcgagagtgag 1767
1377 RRR 1318
1768 tgaagaaggagacacaaaggaagaaagaaactgtagcgagaggttctgcga 1827
1317 RRR 1258
1828 ttctaaccgctgctgagagacaaagaaagatgctactgctgagcgacacttga 1887
1257 RRR 1198
1888 agagattggaatgctgagcctcagattaagagacttgggtagatcctgtagtag 1947
1197 RRR 1138
1948 cacaacccctaagcagagagcatagccatagtgataagcaacctgtgaactggcagccatc 2007
1137 RRR 1078
2008 caagatggaagtcgcaactgctgctgtatttataaagcttaaaagctgagag 2067
1077 RRR 1018

2068 tcta 2071
1017 CGTA 1014

RESULT 2
US-09-094-714A-48
Sequence 48, Application US/09094714A
Patent No. 6117847
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett, Nicholas M. Dean
TITLE OF INVENTION: OLIGONUCLEOTIDES FOR ENHANCED MODULATION OF
PROTEIN KINASE C EXPRESSION
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 6117847rls, LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERECT 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/094,714A
FILING DATE: June 15, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,269
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/478,178
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/089,996
FILING DATE: 09-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/852,852

FILING DATE: 16-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Paul K. Legard
REGISTRATION NUMBER: 38,534
REFERENCE/DOCKET NUMBER: ISIS-2943
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 2244
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-094-714A-48

Query Match 1.0%; Score 47.6; DB 3; Length 2244;
Best Local Similarity 48.5%; Pred. No. 0.0054;
Matches 131; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

2489 ataactatggtttgacaaggttggtaccagttgatagttacatttgggagcaag 2548
511 ATTACCTTAAGGCTGAGGTGCTGATGAAAAAGCTCATGTCACGTACGAGATGCANAA 570
2549 gatcccttcagggaagatgtaggagcgaagaaatcctatgttaagtattactcctt 2608
571 AATCTAATCCCTATGAGATCCAAACGCGCTTTCAGATCCTTATGTAAGTGAACCTTAT 630
2609 ccagatagaagtgtataaataaagaagaagaacaacaacagtcagaagaaacttgaacc 2668
631 CCTGATCCCAAGAAATGAAGCAAGCAAAAAACCAAAACCATCGCTCCACACTAATTCG 690
2669 aaatggaacagacttattattctctcctgcacagagaagaaatccgtgaacgaatg 2728
691 CAGTGAATGAGTCTTACATTCAAATTGAACCTTCAGACAAAGACGAGACGACTGTCT 750
2729 ctggaattacccttgggagatcagtaga 2758
751 GTAGAAATCTGGAGCTGGAGATGCAACACA 780

RESULT 3
US-09-225-749-24
Sequence 24, Application US/09225749
Patent No. 6300320
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
TITLE OF INVENTION: Modulation of c-jun using inhibitors of protein kinase C
FILE REFERENCE: ISIS313
CURRENT APPLICATION NUMBER: US/09/225,749
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 2245
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (28)..(2046)
PUBLICATION INFORMATION:
JOURNAL: Nucleic Acids Res.
VOLUME: 18
ISSUE: 8
PAGES: 2183
DATE: 1990-04-25
DATABASE ACCESSION NUMBER: X52479/Genbank
DATABASE ENTRY DATE: 1993-09-12
US-09-225-749-24

[illegible][illegible]


```

Matches      81;  Conservative      0;  Mismatches      67;  Indels      0;  Gaps      0;

Qy      4410      gctctgtgtgtaaacaccaggttccaagacatcgcgaaccgtatgtcaagtgatctgt      4469
Db      1085      GACATCTGCTTAATCTGATGTGTCCGGACTTTCAGATCCCTATGTCAAGTGAACCTGT      1144
Qy      4470      tagcaagaagggtctgcctagccaaagaagaaccaaaggttgccgaagaagccctggagc      4529
Db      1145      ACCATGCCAAAAGAGANTCTCCAGAAGAAGACTCATGTGAAGAAATGCACCCCAATG      1204

Qy      4530      cccgtaccagaagctctgtccctcga      4557
Db      1205      CAGTGTCAATGACGCTGTTTGTCTTTGA      1232

RESULT      14
US-08-781-891-208/c
: Sequence 208, Application us/08781891
: Patent No. 6090620
:
GENERAL INFORMATION:
: APPLICANT: Fu, Ying-Hui
: APPLICANT: Yu, Chang-Eu
: APPLICANT: Oshima, Junko
: APPLICANT: Mulligan, John T.
: APPLICANT: Schellenberg, Gerald D.
: TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
: NUMBER OF SEQUENCES: 209
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED AND BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
:
COMPUTER READABLE FORM:
:
MEDIUM TYPE: Floppy disk
:
OPERATING SYSTEM: PC-DOS/MS-DOS
:
SOFTWARE: Patentin Release #1.0, Version #1.30
:
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/781,891
: FILING DATE: 27-DEC-1996
:
CLASSIFICATION: 800
:
ATTORNEY/AGENT INFORMATION:
: NAME: NO. 6090620tenburg Ph.D., Carol
: REGISTRATION NUMBER: 39,317
: REFERENCE/DOCKET NUMBER: 240052.419
:
TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
:
INFORMATION FOR SEQ ID NO: 208:
:
SEQUENCE CHARACTERISTICS:
:
LENGTH: 16442 base pairs
:
TYPE: nucleic acid
:
STRANDEDNESS: single
:
TOPOLOGY: linear
:
US-08-781-891-208

Query Match      0.8%; Score 40.6; DB 3; Length 16442;
Best Local Similarity      47.8%; Pred. No. 0.37;
Matches      118; Conservative      0; Mismatches      129; Indels      0; Gaps      0;

Qy      163      gtagtcgtcagaagaagaagagagagagcgtgtgtcctaagatcaagaagaaca      222
Db      16433      GGAGGACGACGAGGCGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCA      16374

Qy      223      caaagacaacacgaacacagtggttccctttagtggatcactgtaactgtaataaagt      282
Db      16373      GGAGCAGAGAGACGACGAGAGGAGAGACGAGACGAGAGGAGACGAGAGAGAGAGCA      16314

Qy      283      tctgcagccccaagaacaaacccaatgtagaaggagccccaagacaagctgcaccaaca      342

```

```

Db   16133 GGAGCAGACGAGCAGGAGCGACGACGAGGAGGAGCAGCAGCAGCAGGAGCAGGAGCAGGAGCAGGA    16254
QY   343 atttgaaatgtataaggaaggcaagtccaagaattgggagaggaaatcgacgcagcagcaag    402
      |         |         |         |         |         |         |         |         |         |
Db   16253 GGAGCAGACGAGGAGGAGGAGCAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGTAGAAGGA    16194
QY   403 gacgaag 409
      |         |         |         |         |         |         |         |         |         |
Db   16193 GAAGAAAG 16187

RESULT 15
US-08-469-802B-6
; Sequence 6, Application US/08469802B
; Patent No. 5741645
; GENERAL INFORMATION:
; APPLICANT: Ori, Harry T.
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Chung, Ming-yi
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene sequence for Spinocerebellar Ataxia
; Patent No. 5741645
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESSES:
ADDRESSSEE: Muelling, Raasch, Gebhardt & Schwappach, P.A.
STREET: 119 No. 5741645th Fourth Street, Suite 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469, 802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Muelling, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110,00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 154 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-6

Query Match          0.8%; Score 39.6; DB 1; Length 154;
Best Local Similarity 57.1%; Pred. No.0.021;
Matches 72: Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY   284 ctctgagcccccacaacaaaaccacgatgagaaggagccccacagaacaactctgccaccaaca    343
      |||||  |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   5 CAGCAGCAGCAGCACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG    64
QY   344 tttagaatgtataaggaaggcaagtccaagaagtgggagaggaatcgcacgcagcagcaag    403
      |         |         |         |         |         |         |         |         |         |
Db   65 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG    124
QY   404 cagaag 409
      |||  |  |
Db   125 CAGCAG 130
```

Fri Nov 23 10:08:47 2001

Search completed: November 21, 2001, 23:59:49
Job time: 10979 sec

us-09-617-099b-2.rni

Db	334	ERRRTRLEKGRSOYSD--RPERDNGRAVED-----	OKORKEEYOTRIRSP	381
Qy	378	NLARVPVPOPEEEMORITHAEVSRARHERHRSDVLANAELEDSRISLIMDRPSROBY		437
Db	382	NLARVPVAPPEEQMRNHNARVSRARHERHRSDVLANPHE-----	AAAAAP	427
Qy	438	SERRAMENORSYSMERTREADOGOSSTPOKRTSNHSEPPTRRSPRPIPLDRPDMRADSLRQ		497
Db	428	AARATGKRAPATARVSPESPARRAAAAAPRPHENGPRPRAPRGAPREPRVPELRQ		487
Qy	498	HLHDPSSAV--RKTREKEMETMLRNDLSLSDPOSESVPRPPRPKSKCKGMQVSLSS		555
Db	488	GRDPGSAVLKARKAREKASLNRDLSLSDPOSESVPRPPRPHRKRGKRRQMSVSS		547
Qy	556	EEELASTPEYTSJCDVLELESSESVESEKOSQKGRKRTSEQVLSDSNTSRBROKRMNYTG		615
Db	548	EEEGVSTPEYTSCEPVELESESVESEKGDLD-----	yy--	579
Qy	616	HLSEEDLMSSEQINDSGVDTSSSTPLNEHSHQKHPTVMOQSKDGRILGRITLILNK		675
Db	560	-----WLDP-----ATWHSNETSPISHPWTQPSKGDRLIGVILNKRT		620
Qy	676	KDGVPYRDSGAMGLKVYGGKMTESGRICAFITTKYKGSJADTVGHLRPGDEVLEWNGRL		735
Db	621	---TWPKESGALLGLKVYGGKMTDGRICAFITTKYKGSJADTVGHLRAGDEVLEWNGRP		677
Qy	736	LOGATPEEYVNIILLESKPRPOVEIIVYSPRIGDIPRIIPDSTHOLESSSSSFEESOKMDRPS		795
Db	678	LPGANNEEYVNIILLESKSPQVEIIVSRPIGDIPIRIPESHPPLESSSSSFEESOKMDRPS		737
Qy	796	ISVTSPMSGMLRDYPOEFLSOGSLIKWTFEDKQHOILVITLGAKDIREDRQGRPRNRYK		855
Db	738	ISVISPISGALKADAPOLVPGOLSVKWLWDKVGHOILVIVLQATDILPRVADGRPRNRYK		797
Qy	856	IYFLPDRDKNKRRRTKYVKTLEPPKNNOFTIYSPVYHREPREMLETILMDQARVREES		915
Db	798	KYFLPDRDOKSKRRRTKYVKTLEPPKNNOFTYVYHVRDRREMLELETIVMDQAPVDOEES		857
Qy	916	EYLGELLILEFTALDDDEPHWYKLOJHDVNSILPRPSPYLPRILOLHGESPTRLQORSK		975
Db	858	EYLGELLILEFTALDDDEPHWYKLOJHDEBSILPRPSPYLPRIHIGESSKKLQORSK		917
Qy	976	ISDSEVSDYDCEDGVGVNS--DYRNGRDLQSLSVPEOVMSNNHSCSPGSPHRADYIG		1033
Db	918	ISDSDISDYEDVDGIGVVPVYGRARABESKATITLVPEOQRTTHNRSRVSPIHRGDDG		977
Qy	1034	PRRSMSPAPRPOURNEOGH--NGTATGUYNTISM--DHRVMDYDHSPPDRDCEA		1088
Db	978	PRRSRLPVN-PIORSDIEHPTRRKSPTTRHNHJASKSPADHRSRHVSQYSSERDESLIM		1036
Qy	1089	ADROPYHRS-----RSTEORPLLERTTTRRSRSERPDNL--MRSMPLMT-----		1132
Db	1037	LPRAKGRSAESELHMTSELQPSLDRKASASTNCLRPDLSLHSPEREHNSKSECSIOKQ		1096
Qy	1133	GNSAPRSPALSRHPRTGSVQTPSPSTPGTGRGRQOLRQLPK-GLTERSAMOLEENRQ		1191
Db	1097	SRKGTASDA-DRTNRCGSPQSPADPSTFSGSRGRQOLRQVYVRSQSTLEOASLVEEBTRQ		1155
Qy	1192	M--KLANKYKOV--AGSDPRLEDOYHKSRYSGMPHRADYVSTSSSDSDVSVANSVTS		1247
Db	1156	KMYKVIHFKQTTGSSGSSQELDHQOYSKYNHNKOYRSCDNAMAKSSSDSDVSVASIBRAS		1215
Qy	1248	SASRESSTIYMSVOSERPRGNRKISVFTSKMORONQVSGSKNLTKSTISIGDMCSLEKD		1307
Db	1216	STRRLSTSSPMOSERPRG--RISSTPRPMQGRMGTSGRAILIKTSVSGEITLERND		1273
Qy	1308	GSGSDPAVAGLIGSGKKRRBSIGAKMAVIYLSKRSKASASOLSTOEGGKKLRETVORST		1367
Db	1274	GSGSDPAVGVGVAGKKRRBSLSKAAVAIY--SKRSSTOLOSOTBEGHKKLAKTIDQST		1331
Qy	1368	ETGIAVEMRMWMTROASRSTDSMNSYSSEGNLIFPGVRLASDSQFSDFLDGLGAPOLV		1427

```

Db      1332   ETGMAAEAMRK -MYRQPSRESTDGSINYSSEGNLIFPEVRVGRDSQFSDFLDGLGPALV    1390
Oy      1428   GROTATPAMGDIQVGMMDKKGGQLEVEITARAGLVAKRGSKTLPAPYKVYLINDNCVIA    1487
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1391   GROTTATPAMGDIQIGEDKKQLLEVIVIRASLTQRKGSKSTPAFYKVYLENGACIA    1450
Oy      1488   KKTTRARKRTLEPYOOLLSEESPOGRVLIIWGDYGRMDHKCFMGVAQTILDELETS    1547
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1451   KKTKTRARTLRDDLQQSLVFDESPPCKVLGYIWMGDYGRMDHKECFMGAQTILLEELDIS    1510
Oy      1548   NMVIGMFKLFPSSLYVDPTSAPLTRRASQSLESSTGTGSTRS    1590
        :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1511   SMIWGMYKLFPPSSLVDPTLAFLTRRASQSLESSSGPCIRS    1553

RESULT      2
T29736
hypothetical protein T10A3.1 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C.Accession: T29736
R.Sanson, J.; Nhan, M.
submitted to the EMBL Data Library, November 1995
A.Description: The sequence of C. elegans cosmids T10A3.
A.Reference number: Z20675
A.Accession: T29736
A.Status: preliminary; translated from GB/EMBL/DDBJ
A.Molecule type: DNA
A.Residues: 1-853 <SAN>
A.Cross-references: EMBL:U41035; PIDN:AAB37028.1; GSPDB:GN00028; CESP:T10A3.1
A.Experimental source: strain Bristol N2; clone T10A3
C.Genetics:
A.Gene: CESP.T10A3.1
A.Map position: X
A.Introns: 41/3; 110/1; 151/2; 194/1; 232/1; 342/3; 390/3; 528/1; 551/2; 628/1; 674/2
```

Query Match	12.1%	Score 1002;	DB 2;	Length 853;
Best Local Similarity	29.7%	Pred. No. 3,5e-42;		
Matches 300;	Conservative 142;	Mismatches 252;	Indels 316;	Gaps 36;
QY 652	HPPTWQSPKCGDRLLIGRIILLNKKRLKDGSVPRDSGMLGLKVVGGKMTESGRLCFAFTIKVK 711			
DB 42	HPPTWQSPKCGDRLLIGRIILLNKKRLKDGSVPRDSGMLGLKVVGGKMTESGRLCFAFTIKVK 96			
QY 712	KGSLADTVGHLRPGDEVLENNGRLLQGFEEVYNIILLESKEPQVEYLVSF---PIGD 767			
DB 97	PGSADTIGRLRPDEVEYENWGQSLQGNATYEQVYISASRYDTSELLYSRAIIIPGCD 156			
QY 768	IPRPDSTHQLQESS-----SSFESQ--KMDRPSI-----SVTSPSPGMV--- 807			
DB 157	--DEFLNLTSPQMSSSAYSRVSAVPSQFORQLPNPDLFLDHPALQOOLSPHQSANFPH 214			
QY 808	-----RDVPQF-----LSQSLIKIMFDKVGHOLIY 833			
DB 215	NNLTLSRNRSTSSYYSDVPLDGVPSNRMEQSQAFGTGHIFGRILEVSEYSHHROLSV 274			
QY 834	TITGAKDLPSDEGDRPRRPYKIYFLPDRSDKNKRTKIYVKLTKPKKMNFTFLYSPVHR 893			
DB 275	ALVNGFLPLRPDSQDTPRNPYKIFLLPDRSEKSRKOSAVIAETLMPVWDEVFYNGITEP 334			
QY 894	EFRMELEITMDARVREESESEFGLIELLETALDEPHMKQLQTHDVSSL--PLPR 951			
DB 335	MLQIRVLELTYWMDKF--GTNSFLGETLIDLASVLDGE-----HSLMCLIVMDMD 384			
QY 952	PSPYLLPRQLHGES---PTRLRQSKRISDSEVSIDCEDGVGVSDYRKNGRDLQSLTL 1008			
DB 385	DNPETRLTKLKKASYNAPTRPQ-----SELNYD----- 414			
QY 1009	SVPEQVWSSNNHCSPSGPHRVDTIGRTREMSPSAPPQPRNVQGHGRGTRATGHNTYSRM 1068			
DB 415	-----HSSNRY----- 420			

QY 1069 DHRVMDHYSDRDRCDAADQPHYRKSSTEDRPRLLERTTNSRSSSERPDTNLMRSM 1128
 ||| :||| :
 Db 421 -----DB-ISON-----IDKOPHHHILAPND-----EENEYIDDELLENDI 456
 QY 1129 SLMTGSAAPSPALSR-----SHPRTSVQTSPPSTPGGR---GR--- 1167
 ||| :||| :
 Db 457 DLATGGGAKKSRTRYRREKGMHGHGVADWTQNHOR---QSGYTSDHGGRQNMIGRAYN 512
 QY 1168 QLPOLPCKTLEKRSADIEERNRQMKLNKKYQVAGSDPRLQDYHSKYRSGMDP---HRG 1224
 ||| :||| :
 Db 513 RROQRPRRSATALISQMEREDM-----YDPRKHRD 542
 QY 1225 ADPTVSTKSSDSVDVSAVSRTSSAFS--STSYMSVQSERPRGNRKISVFTSKMNQNO 1282
 ||| :||| :
 Db 543 DNEYSRSREMTYTN-----RRNMNANTSDNTSF---AETPTANTNVPPIKETQON-- 591
 QY 1283 MGVSCKNLTKRSTISGDMCKLEKNDGQSDPTAVGALGTSKKRRRSIGAKMAIYGLSRK 1342
 ||| :||| :
 Db 592 -----SLASSSSVAG-----GGSANNNVM-----KERKSKLMTFTIPGRABEGR 630
 QY 1343 SRKASQLSQTEGGGKKLRSTVORSTETGL-----AVENRNMNTROASRESTDGSMS 1394
 ||| :||| :
 Db 631 -----RTGFARSEEVGIPGNLSDRLTETPPRLPKQASKSTD----- 668
 QY 1395 YSESEGLIFPGVNLASDSQSPDLGIGPAQVLVGRQTATPAMGDIQYGMMDKGGQLEVE 1454
 ||| :||| :
 Db 669 -SAHSQSNMLP---VLADGRLGTGFVDNIGPGQVGRQVYASPVLGELQIALMALMGRSQIDVE 724
 QY 1455 IIRARGLVYKPGSK-FLPAPYKVVYLLDNGVCVIAKKKTKVARKTEPLPYQQLSPFESPO 1513
 ||| :||| :
 Db 725 IIRAKKLVYKPGSKLFFSAPYKVVYLLMEGKOCIAKAKTNAATKTTISPLFOQLHIFNDSFK 784
 QY 1514 GRVLIIVMGDYGMDHKSFFMGAQIILDELEL--SNMVIQFKLEPPSSSL 1562
 ||| :||| :
 Db 785 KKTLYGTVALGDYGRMERKVFEMGISQIRLEDELGSQPLIGWYKLFHSSSL 834
 RESULT 3
 T00634
 hypothetical protein H_DJ0897G10.1 - human
 C:Species: Homo sapiens (man)
 C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 05-Nov-1999
 C:Accession: T00634
 R:Kaliček, J.; Elliott, G.
 submitted to the EMBL data library, January 1998
 A:Description: The sequence of Homo sapiens PAC clone DJ0897G10.
 A:Reference number: Z114194
 A:Accession: T00634
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-743 <KAL>
 A:Cross-References: EMBL:AC004082; NID:g2822160; PIDN:AAB97937.1; PID:g2822161
 C:Genetics:
 A:Map position: 7q11.23-q21.1
 A:Introns: 36/2; 59/1; 130/3; 172/2; 203/3; 227/3; 288/1; 353/1; 414/1; 461/2; 486/1; 55
 #:119-233/Domains: protein kinase C C2 region homology
 #:Superfamily: protein kinase C C2 region homology <KC2>

Query Match	6.0%	Score 500.5;	DB 2;	Length 743;
Best Local Similarity	23.9%;	Pred. No. 1.9e-17;		
Matches 221;	Conservative 126;	Mismatches 306;	Indels 273;	Gaps 40;
QY	727	EYLENNGRLLQAGTFFEEVYVNIILESKREPPQVELVYSRPIGDPRIIPDSTHQAQLSSSSSF	786	
		:	: : : :	: : : :
Db	2	QYLEWNGIDPLTSKTYEEVOSII--SQSGEAEICVRLDMLMSDENSESHLTHEPPKAV	59	
QY	787	ESQK-----	MDRPSISVTSPMSGML	807
		:		
Db	60	DKAKSGVNDPKQALAEILQKVSLOOSPLVLSVVEKSGVHVSCHSPTSGAGSSVSPSGQPGS-	118	
QY	808	RDVQPELISGOLSTIKLMEFDKGHQILITVILGAKDLPSSREDGGRNRYVKKLYFLPDRSDKNK	867	

Db	119	PSYKRRKGGSKQIIND-LGN-LIHILOARNLVRPNNGXSDPVRKYLLRGCAEYK	176
QY	868	RRKRYAKTLEPKEJNKNQFEIYSVPHRREFRELETTIMDOAVRRESEFLEIIEI-E	926
Db	177	RRKKHVKSLNPFKNQVIYKKSIMEOLKKTLEVTWYDIFR--SSNDFLEVLIDLS	234
QY	927	TALLDEPHYKTL-QTHDVSSLPLRPSRYLPRKQNGES-PTRRLOKSKRISDSEVS	983
Db	235	TSHLDMTPRMYPLKEQESTD-----HGKSHSQSQSQSQPKRS-----	272
QY	984	YDCEDGVGVSDYRHNCRDIOSSTLVPEQYVSSNHCSESGSRHRYDVTIGRRSWSPSAP	1043
Db	273	-----VYKSRHSGIIFPPSKDMQYR-TEKSHSPSGSKSSSE-GHLRSHGPSRS	320
QY	1044	PPQRYNBOGHRGTRATGHYNTISMRDRHVMDHNSDSDRDCSEADRDQRYHNSRSTEGR	1103
Db	321	OSTSTYQTH-----LED-AGAIAAEAAVQO-----	347
QY	1104	PLERTTTRRSRSERPDNLNMRSPSL-MTGKSAPRSPALSRSHPRTGSVOTSPSSPTGT	1162
Db	348	-----LRIOPLAHKSGQS-----NHARKQYRHSIAGVLPIQ	378
QY	1163	GRRGROLPLRPKGTLEBSAMDIEBRNROMKLYKQYAGSPRLEQDYHNSKTRSGMDP-	1221
Db	379	RTOSDNLPR-PANONODS-----QLALRWKMSDGPVKP-----EGAKPP	417
QY	1222	-HNGAD-TVTSKSSDSVSDVSAVSRTSSASFSSTSYMSVQSEPRGNRKISVTFSKMG	1279
Db	418	NHPRAESVSTGSSGSESGSYSVDEGSSSTAGETNLEPT---PRIG-----	463
QY	1280	NROMGVSGMLTKSTISGDMCSLEKNDGSDPTAVGALGTSGKKRRSS-----IGAK	1332
Db	464	---MGQNGQGPVK-----QPRGVGLADTFEARGRGREKPGQACFLGAR	503
QY	1333	MVAIVGLSRKSRSAQSQTEGGKKLRSTV-----QRSTETGLAVERMNMWTRQAS	1384
Db	504	-----NKKETIOPMEPIYETDKDEKYPGSLFLPFSNLCQCIKT-----NM-----	545
QY	1385	RESTDGSNMNYSBSGCL-IFPGV---RLASDSQF-----SPFDLQIGRAQV	1427
Db	546	-KPYD-----EAQADLELPLNLOTRIPLPSMYLLQKTALQYKRSIRGSGNPTINA	599
QY	1428	GROTLATPAMGDIQVGM---MDKKG-QLEVEITIRAGLVVRKGSKT-LPAPYKAVVLDN	1482
Db	600	TTFCEFTTSVMGEIKIALKKEKKTDEQILVELLOCRNITYKRSRPHRLDLYIKIYVMI	659
QY	1483	GV--CIAKKTKYVARKTLEPLYYOOLLSFEESPOGRVLOIIVMGDYGRMDKSTFGVAQIL	1540
Db	660	STQCKVIVKKTTRVCRRHREPSFNETFRSLSPAGHSLOILPNSNGKFMKKTILGECIWI	719
QY	1541	LDELELSNMVIGWFKLEPPSSLVDP	1566
Db	720	LDRVVDLKRRIYVMHKL-----LVSP	740

RESULT 4
T00332
hypothetical protein KIA0559 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00332
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, Y.;
DNA Res. 5, 31-35, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The comp
A:Reference number: Z14086; MUID:98290545
A:Accession: T00332
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1212 <NAG>
A:Cross-references: EMBL:AB01131; NID:g3043641; PIDN:BAA25485.1; PID:g3043642
C:Experimental source: brain
C:Genetics:

A>Note: KIA00559
C:Superfamily: protein kinase C C2 region homology
F:965-1088/Domain: protein kinase C C2 region homology <KC2>

```
Query Match          5.7%; Score 475.5; DB 2; Length 1212;
Best Local Similarity 23.3%; Pred. No. 6e-16;
Matches 213; Conservative 137; Mismatches 297; Indels 267; Gaps 39;

QY 316 RRSQREPOFTEPGLNTRDSRRGRHSKEYIVDEDEVEDVEDEYF---RQREDEYQ- 370
DB |||||-----LADLQSSRLH---SYVKAED-PMEDPEYLKLIKQIKQKQFR 411
DB 363 RRSQEVTFD-----LADLQSSRLH---SYVKAED-PMEDPEYLKLIKQIKQKQFR 411
QY 371 -----ARRSDNLARYPVKPPQPYEQE--MRIAEV---SRAHERR 407
DB 412 GTESLDHLAAGSHYHATSTRHFKSEKYSISRLLLEQAAKQLPAILLYQKOSHKK 471
QY 408 HSDVSLAN-AELEDSSRISLRLMDRPSRQRS-----VSERRAAMENQSYSMERT--- 455
DB 472 LIDPKMSKFSPIQESRD--LEPDYSSYMTSSSTIGTSSRRALQLDDITFGLRKNTDQ 529
QY 456 -----REA-QGOSYTPORTSNHSPPTPRRSP-PLDRPDMRADSL 494
DB 530 QKFMGSSLGTGLTGLNTIRSAIDQDEADKPYSSGSRSPSSSVYGYDL-----SI 582
QY 495 RKQHLDDSSAVRKTRKRMETMLANDSLSDQSSVPRPPRPHKSKGKGKROYSLSS 554
DB 583 KR---DSSSSSLRLKQAEAL---DVSFSAASSARKKPTSLPISQGRGRIPIVAQNS 635
QY 555 SEELAS-----TPREYTSQDVELESESESEKGSQKQKRTK----- 591
DB 636 EEESPLSVGQPMGMARAAGPLPLISADTRQFGSSHLPVQOQIMRESKRTGRYDRD 695
QY 592 -----SEQGVLSDSNTRSEBOKKRMVYGGHSLLEDLEWSEPOIKDSVPTCSITLNEEH 646
DB 696 AFIMDDFOHAMSDEAY-----HLRRETDW-----FKPRESRLENGH 734
QY 647 SHSKHP---TWPQ-SKDGRLI-----GILLNKLRLKQSDVPDSCAMLG 669
DB 735 GLDKRLPRLYHSRLSOHQEOTIQMNGKTMHYTPPHARIKIKTRQSKDHTV---SGNGLS 791
QY 690 LKVVGGKMT--ESGRCLAPITKVKKSLADIVGHLRPGDEVLEWNGRLLOGATFEVEYNT 747
DB 792 IRIYGGKEIPGSHGELIYAKILPFGSAEQGKLMGQVLEWNGIPLTSKYEVEVQSI 851
QY 748 ILESKPEQVELVYSRPIGDIPIRIPDSTHAOLESSSSSEFSQK----- 790
DB 852 I--SQQSGEAEICVRLDNLMSDSSENSQHLLEHPPKAVDKAKSPGVDPKOLAELQKVS 909
QY 791 -----MDRPSISVTSMPSPGM-----LRDVPQ 812
DB 910 LQSQPLVLSVVEKSHVSGPTSAQSSSVSPGQDGPSPVSKKKKSGSKPTDGTQVYSH 969
QY 813 FLSQLSLIKLFDVKQHDLIYVILGAKDLPSREDGRPRNPYVYKIYELPDR-----S 863
DB 970 PITGILQIQIYND--LGN-LIHIILQARNLVPRDNGYSDPFAKVYLLLPGRGQVWVQNS 1027
QY 864 DKKRRRTYVKTLEPKANNQFTIYSPVHRRERREKMLETTITMDQAVYRESEFLEELI 923
DB 1028 AEYKRRTHVQKSLPMEWNOQVYIKYSISMEQLKKTLEVTVMYDHRF--SSNDFLEVL 1085
QY 924 EL-ETALLDDEPHWYKL--QTHDVSSSLPLRPSPLYPRRQIGES--PTRLRLQSKRLSDS 979
DB 1086 DLSSTSHLDNTPKMTPLKEQTESID-----HKSHSQSSQSSQSPSPS-- 1127
QY 980 EYSDYCEDCEGVVSDYRHNHGLQSLTSLVPEQVWNSNHCSFGSPHARVYIGRTSRWS 1039
DB 1128 -----VIKRSRSGIGFPDPSKMQVDP--TEKSHSPSGSSKSSSE--GHLRSHG 1171
QY 1040 PSAPPQKRNVEQGH 1053
DB 1172 PSRSQSKTSYVQTH 1185
```

```
RESULT 5
T25753
hypothetical protein F45E4.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25753
R:Wilson, R.
submitted to the EMBL Data Library, September 1996
A:Description: The sequence of C. elegans cosmid F45E4.
A:Reference number: Z20082
A:Accession: T25753
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1325 <MIL>
A:Cross-references: EMBL:U70852; PIDN:AA09134.1; GSPDB:GN00022; CESP:F45E4.3
A:Experimental source: strain Bristol N2; clone F45E4
C:Genetics:
A:Gene: CESP:F45E4.3
A:Map position: 4
A:Intons: 25/3; 859/1; 928/1; 966/1; 1002/2; 1106/2; 1167/1; 1255/1; 1274/2
```

```
Query Match          5.3%; Score 437; DB 2; Length 1325;
Best Local Similarity 20.2%; Pred. No. 5.4e-14;
Matches 242; Conservative 151; Mismatches 405; Indels 400; Gaps 44;
```

```
QY 7 PRGRPAFTPAASQ-----PPQPEKMDLSHLEERKTIILAYMD 45
DB |||||-----PPQPEKMDLSHLEERKTIILAYMD 45
QY 315 PRYRHNTSVASHSHSTFKHMSKSIACQADELPVYPMISLRQLDMERPPQLMTYVG 374
DB |||||-----PPQPEKMDLSHLEERKTIILAYMD 45
QY 46 RQKKEEKESQVLTKEKHKQPTQWF--PESGITLVNVLQPOKQOPKEKPO----- 98
DB |||||-----LANTESTPTVYRGSTIPRSSREKIASRYRQOQOITYNQMNNND 429
QY 375 KRSKEIQTE-----LANTESTPTVYRGSTIPRSSREKIASRYRQOQOITYNQMNNND 429
DB |||||-----LANTESTPTVYRGSTIPRSSREKIASRYRQOQOITYNQMNNND 429
QY 99 ---TKLHQEFMYKEQYKKMGESQOQOQKQDAPTCGICHTKTFADCGHCNYSQCTK 154
DB |||||-----TKLHQEFMYKEQYKKMGESQOQOQKQDAPTCGICHTKTFADCGHCNYSQCTK 154
QY 430 LLEITK--KYFEDYDQRLREFGERARRHSRRFD-----FHDDDDQ 469
DB |||||-----KYFEDYDQRLREFGERARRHSRRFD-----FHDDDDQ 469
QY 155 FCARCGGVSLRSKVMYVNCVKRQDEILTKSGAMFNSSGNTLQOPDQKVPGLRNEE 214
DB |||||-----FCARCGGVSLRSKVMYVNCVKRQDEILTKSGAMFNSSGNTLQOPDQKVPGLRNEE 214
QY 470 -----DMRKNQVNH--ELARRKRCAS----- 490
DB |||||-----DMRKNQVNH--ELARRKRCAS----- 490
QY 215 APQEKAKLHQPOFGAPGDLSPVAVEKGRANGLTRDITKNGSGVKGQLASDMPSDRK 274
DB |||||-----APQEKAKLHQPOFGAPGDLSPVAVEKGRANGLTRDITKNGSGVKGQLASDMPSDRK 274
QY 491 -----CEIISTDPRTA----- 503
DB |||||-----CEIISTDPRTA----- 503
QY 275 RSPSVSRDQNRRYEQSEREDYSQVPSDGTMPR-----SPSDYAD--RRSQREPOEYE 326
DB |||||-----RSPSVSRDQNRRYEQSEREDYSQVPSDGTMPR-----SPSDYAD--RRSQREPOEYE 326
QY 504 LNPVVPMS-----SDYSHVPHYGSLPRIDYPRGTRSDVRDPTVYRQHLPOQPP 552
DB |||||-----SDYSHVPHYGSLPRIDYPRGTRSDVRDPTVYRQHLPOQPP 552
QY 327 EPGLNTRDSRRGRHSKEY-IVDEDEVEDSEDEKORREDEYQARR--SDNRLARY 382
DB |||||-----EPGLNTRDSRRGRHSKEY-IVDEDEVEDSEDEKORREDEYQARR--SDNRLARY 382
QY 553 TNYSYNAGSLPRNFERGFSDLPIETENEGSFGAPRRHRSNLGYESTSMFNLSDEVYLG 612
DB |||||-----TNYSYNAGSLPRNFERGFSDLPIETENEGSFGAPRRHRSNLGYESTSMFNLSDEVYLG 612
QY 383 PVKQPOEYEQMRIAEEV-----SRAHERRHSDV-----SLAN-----A 416
DB |||||-----PVKQPOEYEQMRIAEEV-----SRAHERRHSDV-----SLAN-----A 416
QY 613 DSIPIVQHQDPRVYDQIPRSGYAQDTNLNLSNGVYKSGDMYSQVSLNSQPOSLQOSA 672
DB |||||-----DSIPIVQHQDPRVYDQIPRSGYAQDTNLNLSNGVYKSGDMYSQVSLNSQPOSLQOSA 672
QY 417 ELEDSSRISLMDRPSRQRSYSEERAAE-----NORSYSMERTREAOQSSSYPTQSN 470
DB |||||-----ELEDSSRISLMDRPSRQRSYSEERAAE-----NORSYSMERTREAOQSSSYPTQSN 470
QY 673 QLPQPMIPITRYDAPMSDPYVSSRSRLVDQNMQNGVNSQVYHNHSTLLAPKSAQ-----MT 727
DB |||||-----QLPQPMIPITRYDAPMSDPYVSSRSRLVDQNMQNGVNSQVYHNHSTLLAPKSAQ-----MT 727
QY 471 HSPPT---PRRSPILPDRPMPRRADSLRKQHNLDPSAAVAKTKREKMETMLRNDLSSD 526
DB |||||-----HSPPT---PRRSPILPDRPMPRRADSLRKQHNLDPSAAVAKTKREKMETMLRNDLSSD 526
QY 728 HIAPTYNQLAQOQGLIPMONTQM--DPLMSSGRISSSSQVYSNEM----- 772
DB |||||-----HIAPTYNQLAQOQGLIPMONTQM--DPLMSSGRISSSSQVYSNEM----- 772
QY 527 QSESVRPPRPHKSKGKGKROYSLSSSEELASTPYTSCDVELSESVSEKGDQK 586
DB |||||-----QSESVRPPRPHKSKGKGKROYSLSSSEELASTPYTSCDVELSESVSEKGDQK 586
QY 773 -----NGSRPAQSSLF-----EYGNRQYGAAPPTTYDVPNVYASAD--- 810
DB |||||-----NGSRPAQSSLF-----EYGNRQYGAAPPTTYDVPNVYASAD--- 810
QY 587 GKRTSBOGVLSDSNTRSEBOKKRMVYGGHSLLEDLEWSEPOIKDSGVPTCS----- 638
DB |||||-----GKRTSBOGVLSDSNTRSEBOKKRMVYGGHSLLEDLEWSEPOIKDSGVPTCS----- 638
```

Db 811 -WRITVOGSLIQMPKMOQOQHNTF-----DARWPRDLSLRKATATSRRAQETA 861

QY 639 -----STLNEEHS--DKHPTWQPSKDDRLIGILLNKL 675

Db 862 LTTSMKISTGSRSYARRPIRPSRYRNPDEATNSMDRH--VARRTAEANSRYDKRILLTFTSY 920

QY 676 KQGSVPKDSGANLGLKVVGKMTESGRLCFTIKVKKGSADLYGHLRPGDEVLEWGR 735

Db 921 KHNHNYND---LGVNVGSGKROMNGELSAVVSQLSHTANNQTLGQIKIGDEVLEWNGIL 976

QY 736 LGATFEEYVNIJLEKPRPVELVVSRIQDIPRIIPDSTHAOLESSSSPESQMDRPS 795

Db 977 LRGKTFEEYERIV--KNSGEIEMVI-----RTYKPPSSGYVDITPLRNK- 1019

QY 796 ISVTSPMSFGMLRDYPOF-----LSGQLSIKLMEFDKVGQHLIVTI 835

Db 1020 -TMRDRLSDPRVPRVPMHNRINGINNNSVLNHNHTLSDSCCHIGIUVSLGUDQ-NSRLVAKI 1077

QY 836 LGAKDLPSBEDGR--NPVYKIGFLPRDSKKNKRRRTYVKTLEPRKNQPTIYSVNHKE 894

Db 1078 IRRGKLSKSDQSSADNPKEKVVLLPRGKVSHKRTKTRFVSSCAPENWQVLEYGQ-VADPT 1136

QY 895 FREEMLEITLMDQARYREESEFLGELIIL-ETALLDDEPHNKLQTHDVSSLPLRPRS 953

Db 1137 LMTMFLFTYCDVQ--RDVDDPLGYNQIIRLADKSAINTGRWVPLQ-----GSDQRP 1190

QY 954 PVLPRQLHGESFTRRLQRSKRISDSEVSDYDCEDGVGVVSDYRHNCRDLOSSTLSEBQ 1013

Db 1191 HYM-----NGTSLS-----QIPSI 1204

QY 1014 WMSNHCSPSGSPHRYDVIGRTSMSPSAPRQNRVBOGHGTR-----ATGHYN 1063

Db 1205 AASKHSAVAHMH-----NHNISEVPSTILYPKGVGTRHDPKRVNRATFNYN 1253

RESULT 6

T34318

hypothetical protein K03A1.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34318

R:Nhan, M.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid K03A1.

A:Reference number: Z21505

A:Accession: T34318

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-547 <NHA>

A:Cross-references: EMBL:U41625; PIDN:AAA83324.1; CESP:K03A1.3

C:Genetics:

A:Gene: CESP:K03A1.3

A:Introns: 33/3; 50/2; 95/1; 118/3; 181/2; 293/3; 342/1; 371/3; 390/2; 451/1; 484/2

```

Query Match      4.38; Score 358; DB 2; Length 547;
Best Local Similarity 23.08; Prid. No. 1,4e-10;
Matches 140; Conservative 99; Mismatches 225; Indels 146; Gaps

QY 33 TEERKILLAMVDQKKEEKESQVLKKEBKAPQWMPFSGITELVNVNVLQPOKOP 92
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 TONNGYAIKRYFKKROKBEAKE-----TQ-----ISKASEELSDIKOI 63

QY 93 NEKEPOTFLHQGFEMKYKQYKMKGEESQOQOEOQGDAPPTGICHTKTFADGCGHNCSYCO 152
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 TERKETSK-----KLVTQ-----DPAIQICQKTFADGIGHKCFYCO 102

QY 153 TKFCARCGGRVSLRSNKVVMVVCNLCRQOEILTKSGAMFYNSGSSNTLQOPDQKVPKGLRN 212
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 LRSARCRCGRQSK-NNAIWMCSLCQOROOILAKTKGMF-----QP----- 142

QY 213 EEAPEQKAKLHEQPOFGAGAGDLSVAVEKGRAGHLTROTDTIK-----NGSGVK--H 263
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]

RESULT 7
S44644
Hypothetical protein F37A4.7 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Sep-1999
C:Accession: S44644; T16287
R:Fulton, L.
submitted to the EMBL Data Library, February 1994
A:Description: The sequence of C. elegans cosmid F37A4.
A:Reference number: S44638
A:Accession: S44644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1021 <FULT>
A:Cross-references: EMBL:U00032; NID:9458958; PID:9458965; PIDN:AAA50635.1; CESP:F37A4
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F37A4.7
A:Introns: 74/1; 101/2; 153/2; 198/3; 234/3; 344/2; 462/1; 519/1; 592/3; 616/1; 741/1
C:Superfamily: protein kinase C C2 region homology
F:736-851/Domain: protein kinase C C2 region homology <KC2A>
F:876-991/Domain: protein kinase C C2 region homology <KC2B>

```

Query Match          4 1%: Score 358; DB 2; Length 1021;
Best Local Similarity 20.9%; Pzed. No. 3,1e-09;
Matches 197; Conservative 139; Mismatches 342; Indels 266; Gaps 43.

QY 90 KOPNEKPEOTKTHQOFEMKFEQVKKM---GESSQOOOEQR-----GDAP 130
      : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 114 RSPNTSKAOTSITAAE--QEHLOKVLAAEKSEKSEQRIGKMDRLKMRPRATNGV 171
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 T-CGICHKRFK---ADGGCHNCVCOOTFCFAR-CGGR---VSLRSNKV--MMWVCLCKR 179
      : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 172 TCHCLCH-TEFGLLSKSTAAACVCDRKVYVCOBNCGEVETTDYNOTTGKYEYFLCLTCE 230
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 QOEIL-TRSGAFNYSNGSFTLQOPDKVPRGL--RNEEAPQEKRAKLIIEOPQFGAPG 234
      : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 AREVLIMKKSGAMFYKE-MPEFQRPDDRLRYVYPVNTGLPMASSA---APPLSTPGG 285
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 ---DLSVAVE-----KGRHAGITTDQTLKNGSGVYHQIADSPDRKRS----- 276
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
Db 286 AGPDMTPSTSSCOMTTPKMASPGVNSPGLQMGNP-----TSPLENGTRRRTGEGIE 341
OY 277 -PSVERDQRRARRRQSEEREDYSQYVP-----SDGTPRSPEDYADRRSQRRQFEEEGHL 331
Db 342 PRSSSDGSEVQSGVPRALNKKTPVGSTSATTSAPRPTSTTPTRRRE-----A 392
OY 332 NYRDSNRGRHRSKEYIVDEDEVESRDEYERORREERYQARYSDPNLARYPYKPOPEE 391
Db 393 NMERRSRHHAHANRLYSTVDDDDSSPESRPTSTRS-----PRHSL----- 434
OY 392 QMRTHAEVRAHRRHSDVSLANAELEDSRISLILMRDRSRQRSVERAAMNQSY 451
Db 435 -----ATPSYADTCH-DTSLPDADTRSIDSGVQSDHSPQOS-----GLTCS 478
OY 452 MERTPAQOOSVPORTSNHS-PPTRRSPILPLDRPMRADSLRKHOLDPSSAVRKT 510
Db 479 SSSLPLQQAASH-----DHHSGGTERR-----ISNP-----RTS 511
OY 511 REKMETMLRNDLSLSDQSEVPRPPRPHKSKGKKRQVSLSSSEELASTPEYTSDD 570
Db 512 RV-----AQSAGTSLVTPPP-----ISSRTPDNQNSSPLNV----- 545
OY 571 VELSESYSKGSQCKRKT-SEQVGLSDSNTRSEROKKMYGGHLEEDLEKSEPOI 629
Db 546 --MEKSSSASTASSGGRVGSAPVLNHMHAMNQ-----NHNDIKKLISQTSR 596
OY 630 KDSGVDTCSSTTNEHSHDKHPVTWQPSKDGRLIGRILNKLKDGSVPRDSGAMIG 689
Db 597 AESPLAASSSFLSPDDDTQKN-----RRRDYGVRVNSLQRTSLDDVAPP----- 643
OY 690 LKVVGGKMTESGRCLATTVYKKGSLADYGHILRPGDEVLEWNGRLLQGAFFEEYNTIL 749
Db 644 -----VAPISKMN-----GH-----IV 655
OY 750 ESKPEPQV---ELVSPRIDIPRIDSTH--AOLESSESSFEQKMDRPSISVTPSPSP 804
Db 656 SSEPSTSTNOMHTSVPIPTVYVPEEEKALITASTESSEPGAVIEPLDIDENIEP 715
OY 805 -----GMLRDV-----PQFLSGQISIKLMPDKVGHQILVITLGAKDLPSREDG 847
Db 716 KHASRRRGDVGVRGFCILCFSRKQRSIGSTILTLTYHSAKCLKMLHILIRKKNLKMAMSN 775
OY 848 RPRNPYKATIFLDRSRKKNRRTKTKVKKLEPKNNOPTISPVNHRREPRERLETLMDQ 907
Db 776 GFSDPYKFLPLPENTATYKLTSTKTEKTLNPEWMEHMYSTGITEDDEKKILRVTVD 835
OY 908 ARVREESEFGEILIELETALLDDEPRHMYKLQTHDVSSLPPLR 951
Db 836 DRI---GSDFLGERTIALK-KLNDENMKKFNLYLE--SALPYQ 873

RESULT 8
A45973
trichohyalin - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 01-Mar-1996 #text_change 22-Jun-1999
C:Accession: A45973
R:Lee, S.C.; Kim, I.G.; Marekov, L.N.; O'Keefe, E.J.; Parry, D.A.D.; Steinert, P.M.
J. Biol. Chem. 268, 12164-12176, 1993
A:Title: The structure of human trichohyalin. Potential multiple roles as a functional E
ed (cross-linking) protein.
A:Reference number: A45973; MUID:93280194
A:Accession: A45973
A:Molecule type: DNA
A:Residues: 1-1898 <LEE>
A:Cross-references: GB:LO9190; NID:9292835; PIDN:AAA5582.1; PID:9292836
A:Note: authors translated the codon AGG for residue 1714 as Pro
C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sh
Covalent modifications to this protein include conversion of arginine to citrulline and
C:Genetics:
A:Gene: GDB:THH
A:Cross-references: GDB:136223; OMIM:190370
A:Map position: 1q21-1q21
```

```
C:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
F:49-81/Domain: calmodulin repeat homology <EP2>

Query Match 3.7%; Score 307; DB 1; Length 1898;
Best local similarity 17.4%; Pred. No. 2,3e-07;
Matches 245; Conservative 213; Mismatches 514; Indels 434; Gaps 53;

OY 19 QPPPEMPDLSHLEEE--RKTLAVMDRQKEEEKESVLYKKEHKAQPIQWTFPSG 76
Db 264 EPQORELOE-----EEQRLKLEEROLELRERQEEEOOQRLRREOQLRKDE----- 312
OY 77 ITELNVNVLQPOQKQPNKEPEQTKLQOFEMYKEQYKKGSESOQOQ---EQGDAPTCG 133
Db 313 -----ERREQOERREQOERREQOERREQOERREQOERREQOERREQOERRE 358
OY 134 ICHTKFADCGCHNSYCQTKFCARCGHVSLSNKMVMVNCILRQOELITKSGAMFYN 193
Db 359 -----REOQL----- 364
OY 194 SGSNTLQOPDQVPRGLNBEARQOEKKAHLHEOPQOGABDLSVAVEGRAHGLTRQD 253
Db 365 ---RREDEERREOQLRREOERREOQLRREOQLRREO-----QURREOQLRREQ 412
OY 254 TIKNSGVKHOLASDMPDRKRSPSVSDONRRYEQSEER-EDYSQYVPDGTMPSPSD 312
Db 413 QLRREOQLRREOQLRREOQLRREOQLRREOQLRREOEEHQBKHDERREOQLRKROEE 472
OY 313 YAD-----RSQREPOFEEEPGH--LNYRDSNRGRHRSKEYIVDEDEVE 355
Db 473 RBDWLKREETERHNEQERQOQLRKQOEEERRERWMLKEEERREOERREOQLRROEE 532
OY 356 SRDE-YERQREERYQARYSDPNLARYPYKQPYEROMTHAEVSARHRRHSDVSLA 414
Db 533 RREOQLRKQOEEERLEOQLRSEQOQLR--EQEBRLQQLKREEEKYLEDQRRQRL-- 586
OY 415 NAELEDSRISLMDRPSRQSVSERRAAMENQSYSEMTREAGGSSVTPQTSNHSPP 474
Db 587 KRQOERROQLKREERERROQLRKQOEBRLQQR-LKREVEYLEQERBERDLKREPE 645
OY 475 TPRRSPILDRPMRADSLRKQHLLDPSSAVRKTKEKKETMLRNDLSLSDQSESRPP 534
Db 646 EERRH-----DLKSEDEERRH--DQLREOERREOQLRRE--EEERLEQR 690
OY 535 PRPRHKKKGKKNQVSLSSSEELA-----SPEYTSDDVLESESESEKQDSQK 588
Db 691 LKREHEER---REOFLAEEEOQANERIKSRIPKM---QWQLESEA-----DARQSK 737
OY 589 -RKTSEQVGLSDSNTRSEROKKMYUGHSLDEDEWSEIQIKDSGVDTGSSTLNEBHS 647
Db 738 VLEAPQAGRAEAPQOEBEKKRR-----ESELQOE-----EERRA 772
OY 648 H-----SKHPVTWQPSKGDRLIGRILNKLKDKDSVPRDSGAMIGKVVGKMTESG 701
Db 773 HRQOEEEOQRDRFTWQAEKSEKRGRLSAR-----PRLEQEREROLEEEOQRQR 827
OY 702 RLCAFLTKVKKSLADYGHILRPGDEVLEWNGRLLQGAFFEEYNTILLESKPEQVELV 761
Db 828 FL-----PEEKEQQR-----QRRREKELQF 851
OY 762 SRPGLDIPRIDSTHAQLESSESSFEQKMDRPSISVTPSPKMLRDVPLFSGQLSK 821
Db 852 -----LEE-----EPQLRERRAQOQLEEEEDQOEEEOERROQRQDQ 889
OY 822 LMPDKVGHQILVITLGAKDLPSREDGPRNRYVYILFLPDSDNKRRRTKTVKTLPEK 881
Db 890 KW-----RWOL-----EEKKRRRHT-LYAKPALQOQLRKQOQLQEEEBELQ 931
OY 882 NQPTISPVNHRREPRERMLETLMDQARVREESEFGEILIELETALLDDEPHMY--- 937
Db 932 REE---REKRRRQOENQYR---EEQLOQEBEDQLRE---ERKRRRQOERERQYRDX 981
```

```

Qy 938 KIQTIVSSLPRLPRSPYPLRROLHGESP--TRRLQBSKRISIDSEFYSDVDCEDGVGVSD 995
      |||
Db 982 KLOOKE-----EQLLGEPEKRRRREKREKYYREE-----1011
Qy 996 YAHNGDLQSSLTVEPQVSSNNHCSPSGSPHRYDVIGRTFSMS--PSAPPRQNVQEG 1052
      :||
Db 1012 -----ELOE--EQLLREEREK-----RRQEMEROYRKKDELQEEEL 1050
Qy 1053 HNGTRATGHYNTISRMDRHHVMDHYSSDRDRDCEAADRQPHYRSTSEQRPLLETTTR 1112
      |
Db 1051 LKEER-----EKRRLOEREROYREEEELQEEELGEEERETRRQLEEROYRK 1099
Qy 1113 SSSSEPRDNLNMSMSPLMTGRSAPSPALSRSHPRTSVQTSFSPSTPGTGRGRQLPOL 1172
      :||
Db 1100 EEELQEEELLE-----EPKRRROEREROCREEEELQEEELREEREK 1147
Qy 1173 PEKGTLEERSAMDIERNRQMLNKYKQVAGSDPRLQODYHSKYSRSGMDH-1222
      :|||
Db 1148 RRRQELEROYREEEELQORRKOQYR-----DEQORSLLKQWEPKEKENVADNKV 1198
Qy 1223 --RGADTVSTKS--SDSDVSDVSAVSRTSASRFSSTYSMSVOSEREPRGNRKISVFTSKMQ 1279
      :||
Db 1199 YCKGRNEQFROLEDQSVDR-----RQSQQDQLHLLGEEQENDR-----EQE 1240
Qy 1280 NRMQVSGKNLTKRSTISGDMCSLEKNDGOSDPAVGAIGTSGKKRRSSIGAKMAIVGL 1339
      |||
Db 1241 RRRWQANRHFPPEE-----OLEREBQKEA-----KRR-----1268
Qy 1340 SRKRSASOLSOTEGGKKLRSTYQR 1365
Db 1269 DRKSQEEKQLREEREKRRRQETDR 1294

RESULT 9
T49316
Profileagrin related protein [imported] - Neurospora crassa (fragment)
N:Alternate names: protein B13N20.10
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49316
R:Schulte, U.; Algrn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49316
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1386 <SCH>
A:Cross-references: EMBL:AL355925; GSPDB:GN00116; NCSP:B13N20.10
A:Experimental source: BAC clone B13N20; strain OR74A
C:Genetics:
A:Gene: NCSP:B13N20.10
A:Map position: 6

Query Match 3.4%; Score 278.5; DB 2; Length 1386;
Best Local Similarity 20.7%; Pred. No. 4e-06;
Matches 284; Conservative 163; Mismatches 540; Indels 385; Gaps 62;

Cy 89 OKOPNKEPQTKLHQEFEMTKEOYKKNKGESQOQEQKGDAPTCIGICHKTKFDGCGHNC 148
      :|||
Cb 58 EFEREREREPKARFESYE-----VETVVEEREEREIYGG-----GDARRS 97
Gy 149 SYCQTFQFCRCGGRGVSLSRNKVMVWVNCNCRQOELLTKSGAFYNGSGNTLQDPQ---- 204
      |
Db 98 STRESSQSSRRHH-----GRDGHRRSSHRSSEEREWT 132
Qy 205 KYPRGRL--NEEAPQOEKKALHE--OPQFQAPGDLISYPAVEKGRAGHGLTRDITIKNS 259
      :|||
Db 133 RERGRGRETTELVEEYREKNEYHESEHYILGGRHSSGGEVEYSYR-HGSSRDRDSQOK 191
Qy 260 GYKHQIASMDSPDRKSPS--VSRDQNRREYQ-----SEEREDYSQVPSDGTMPR 308
      :|||
Db 192 A-----SSSRRRSEERETREDRERREQGARVTEELTEVEEYREYH----- 232

```

QY	309	SPDYADRRSQRPPQYEEBCHLNTYD-----SNRGHSHSEYIYDDDEVEDRD----	358
Db	233	-----BSEHRCMGHGSESRHSHSHNSRDRGKASSSSGGHGHKSSGDOERAREVETEE	291
QY	359	-----EYEROREEEOAYRSDPNLARYPVKPOPYEOMRIAEYSRABERHRSDVSLANA	416
Db	292	-----IYEEERHAYEEERGGRRS-----HYEESH--RHWKRGPGS----	326
QY	417	ELEDSHISLRMD---RPSQRBSVSERRAMENORSTSMRTENAOG---QSTPQSTN	470
Db	327	---QSSRASDSKODGHTRGSKERETRETEADEDRHGEERREGORLGEVESS---ORQES	382
QY	471	HSPPTRRSPRIPLRDPRMDRMLARKHONHDPSSAVKTKREK-----M	514
Db	383	HSD-----KGGDSWRDYEKSSSTPEPKRGAELEARDPRAEEOGKHQASRGL	427
QY	515	ETMLRNDLSL---SSDOSESVRRPPRRPKSKKG--GKMRQVSLSSSEELASTPEYTS	568
Db	428	EASRORESLGRKTKSSSDOLR-----GLQESGATEVERARHNFTGEYEYDT	476
QY	569	DVELESESVSEKGDQKGRKTSQ---GYLSDNSTREOQKKRMYUGSHLEEDLE-	623
Db	477	PAAE--ESHOREESPTEHGRGHEHEDYOANCMBEEBCHYQAOQGEOMQONEEEEEEE	535
QY	624	-----WSEQIKRDSGVDTCSTTLNEE-----HSHSDKHAPYTWOPSKD--GDR	664
Db	536	EDLEFLEAKRRSPAVK-----AKSTLKRNDPASFLEPHQOPQSSRHMHSNKQAWRDQ	588
QY	665	LIGRIILNKRKKGQSVPRDSGAML---GLKVGGKMTESGRLCATITVKKGSGLADTVGH	721
Db	589	LMTADKSELEFEESLHSHSRAKFRPLRPYVNKKANLSSFESSPAPRDEISLQEGBN	648
QY	722	---LRPDEVL---ENWGRLOGATEEVEYVNIIESKREPOVELVAVRPIDIPRIST	775
Db	649	ASLRPPEVUDROGQANGKVSOGSGLSEKSOHLPTQOEABRKIGTT-PLG---RVSSQ	704
QY	776	HAOLESSSSSEF-----SOKMDPSISV---TSPMSGMLDVP	811
Db	705	RUSVPYSPKPFQPSRIGTAPVLGSGSSRPLOPQORVLRHSSIIIGSSSPGSGIYPPV	764
QY	812	QPLSGQLSTIKLMDKYGHOLIVITLAKDLPNSE--DGRPNRPYKIKFYFLDR-----	862
Db	765	Q-----KONTTPARK-----PKAK-PLERERDGGTGTVTALRPLDNHSSGGGTT	810
QY	863	---SDKNKRRTYVKTKTELPKN---QTFIYSPVAREFEREMELITLMDQARVEEE	914
Db	811	PNSTPKRQOLRGLMESIKSLRSDPRKAVKEGRENENE-----EEOGREED	858
QY	915	SEFLGELIELEFALLDDERHNWKLOTHNVSLRLPRSPYLR--RQOLNGESPTTRIQ-	971
Db	859	MODYDEI--VELEEQSSAAE---SFERRPLRSLRDKNNSPISKALSALRPRTKGTGRGAVO	914
QY	972	RSKRISDSVSVSDCEDGCVGVVDYHNHNRDQSLTSLVPEQWMSNHCSPSSPHRYD	1030
Db	915	ANSVROSPTRRNADVEQRRV---EGGR-----STRVITTSSTPTHTESP----	958
QY	1031	VIGTRSWSPSAPRRORNVBOGHRGTATGNHNTISHMDRHNVADHYSSDRDRDCEAAD	1090
Db	959	-----VHSPAGSEBIE-----NATGMEDYDQDEQNEHEDTVDMDMDVMD-D	1003
QY	1091	KOPYHNS-----RSTEOPLLEKTTTJRSSSRRTNLMKSRMSLTGCSAPRS--PA	1141
Db	1004	LEMVESHSPAPRAVQAPRRAPRRPRLSPRGKROQOQOQOELQOQVAAARORPRAOR	1063
QY	1142	LSRSHRTGVSQTPS-----STPGTGRRGROLRQOLRPGK-----TLERSAMDIEER	1188
Db	1064	LKRP---TSTLEMSPYLRSGVSLFEGYGRKRSLLRPRLPRSNMNIYTPRLQPLRSLPVSSF	1120
QY	1189	NROMKLINKYQAVASDPRLBODYHYSKTRSGMDHRG--ADTVSTKSSSDSDVSDVASVST	1247
Db	1121	GOO---QOQOQOQOQSSPSTSTSON-----HNGCAGTSGASIGASSSTNSASAAATTTT	1168

QY	1248	SAS-----	-----RFSSTYSNVGSQERPGANKISVFTFSKMQRMG	1284
Db	1169	SHSHDHHQOOOOLHPRLPTLHMSTSHMDLSTLIHNAHRLHNSQSOSQSGKG	1220	
RESULT	10			
	113564			
	microtubule-associated protein homolog - fruit fly (Drosophila melanogaster)			
	N:Alternate names: hypothetical protein EG:49EA.1			
	C:Species: Drosophila melanogaster			
	C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000			
	C:Accession: T13564			
	R:Spanos, L., Papadogiannakis, G.; Siden-Kiamos, I.; Louis, C.			
	A:Submitted to the EMBL Data Library, April 1999			
	A:Description: Sequencing the distal X chromosome of Drosophila melanogaster			
	A:Reference number: 217689			
	A:Accession: T13564			
	A:Status: preliminary; translated from GB/EMBL/DBJ			
	A:Molecule type: DNA			
	A:Residues: 1-5327 <SPDA>			
	A:Cross-references: EMBL:AL031128; PIDN:CAA20006.1			
	C:Genetics:			
	A:Cross-references: FlyBase:FBgn0025392			
	A:Introns: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1			
	A:Note: EG:49EA.1			
	C:Superfamily: Drosophila 576k microtubule-associated protein homolog			
Query Match	3.3%;	Score 272.5;	DB 2;	Length 5327;
Best Local Similarity	18.5%;	Pred. No. 4.3e-05;		
Matches	278;	Conservative 216;	Mismatches 575;	Indels 431; Gaps
QY	24	PEMPDLSHLTERRKILLAVMDROKKEEKEOSVLKI-----	60	
Db	1066	PLTFEDERIRPLDQIKEDLVIEEKYVEETKEAEIVATVQTLPEAPLAIDITILASATK	1125	
QY	61	--KEEKAQPTQWFPSSG-----ITLVNVNVLQPOOKNEKPEQTKHQEFM--	107	
Db	1126	DAEKDANNAELGELPSGGERVLPMAKTFEAOONLKDVIKTDE--VADLPVINEADLGIX	1184	
QY	108	-----YKQOVKMKGEESQOOOEOK-----GDAPT--CGICHTKTFADSCGHN	147	
Db	1185	EKMSQDANKSIHKEESAKKEKETDDEKENKGEIETLDEPKVDISH-----	1233	
QY	148	CSYCTKTFCAKCGGRSLR-----SNKVMVNVNLCRKOELITKSGAMFYNGSNTLOO	201	
Db	1234	-----VLKESVOEAKVAVITTYEKKOEELVEA-----TTVITQ	1270	
QY	202	PDQKVRPGLNEPAPEKKAHLEQPOFOGAPGDLSPVAYEKGRAGLFRQDTIKNGSV	261	
Db	1271	ENQE-----DLMEQVKDKEEHQO-----KIESGI-----ITEEAKKSASTP	1307	
QY	262	KHQIADMPEDRKRSPEVSQDQNRRIQSEERDYSQVYSDDCTMPRSP-----DY	313	
Db	1308	EKEKTSIDTSDDELPAQLAPPTTVPPKSAKDRED-----TQSIESPITTEAIEVQ	1360	
QY	314	ADRESOR-----EPQFYEEPGHLNYPDSNR--RGHRHSKEYIYDDEVEDSEDEYENQRE	366	
Db	1361	AKQDAQKVPVAREAIKTESPLASKTSPESATGSYKE--DTEQTKSKKSPVSRPE	1417	
QY	367	EEOYARVRSQDNLARVYKQPOPYEOMKRIAEVSRANHERHSHVSLANMLEDSTRSL	426	
Db	1418	SE--AKDKSPFAGSGAESRPESVASEYKDEGKAESRE-----STAKTHKDS	1464	
QY	427	RMQDRPQRQSRVSEERRAMENQSRYSMTREPAQOSSYVPORTSNHSPPTRRSPIDLRP	486	
Db	1465	SLDPAKDESRRESIAESIKFESGIDERSALASKEASRPEVDTQKSPRESREI-----	1519	
QY	487	DMRADSLRKQHLND-----PSSAVR--KTKREKMETMLKNDLSLSSQSSSVR	532	
Db	1520	-----AESLKAJSTKDEKAPPSKEASRPQSVASEYKDETEKSKSPRSRESI--AESAK	1571	
QY	533	PPRH-----KSKKGKMKRQYSLSSSEELASTPE--YTSCDQVLELS--	575	

Db	1572	PIPIEFREVSRPEESVIDGIDKDESAPESRRDPSPLSKSE----	ASRPESYLESEKDEPIRISTE	1628
QY	576	-----EVSVE--KGOSOKGRKRTSEOGVLSDSNSTRSEROKRMYYGGHSEEDLEWSEPQI		629
Db	1629	KSRESVAESKADSTKDEKSP-----LTSKDISRPEASAVENM-----	DAPFKETSR	1676
QY	630	KDSGVDFCSSTTLNEEHSKSHKPYTWOPSKDQDRLLGRILLNRKDKGSVPRDSGAALG		689
Db	1677	PESAVGSKMKDESMSE-----PSR-----RESVKGA-----		1703
QY	690	LKVVGKMTTEGRCAPITTKVKKGLADTVQHL--RPGDEVLEWNGRLLOGATFEEVYNI		748
Db	1704	-----AQSRETSRPASVAESAQD--ADDLKELSR-----ESTYOSKEAGSI		1744
QY	749	LESKEPEOVELVSRPIDIPRIPIDSTHAQLSESSSSFEQKMDRPSISVSPMSPGLMR		808
Db	1745	KDEK-SPLASEFASRPASVAESVKDEAEKSKESRRESVAEKSPLPSKESAPASVAE--		1801
QY	809	DVPOPLSQLSTIKILMFEDVYGHDLIYITLIGAKD-LPSREDGRPRNRYVYIYLPIDRSKRNK		867
Db	1802	-----SIKDAEAKSKESRRESVAEKSPPLPSKESRPASVAESIKDAEAKSKES		1851
QY	868	RRTTKTVKKTLEPKMNOQTFLYSPVHRR-FREEMLEITLMDQARVEEES--FELICEILIE		924
Db	1852	RESVAEK-----SPLPSKESRPASVAESIKDAEAKSKESRRESVAE----		1895
QY	925	LETALLDDEPHYKTIQTHVSLPLPRSPYLPRLQHLG--SAPTRLQRSKRISDSEVS		982
Db	1896	-----KSLPSKASRPASVAESIKDAEAKSKESRRESVAEKS		1934
QY	983	DVDCEDGVGVSDYRRHNGDLQSLTVPEQYWMSSNHGSPSGSPHRYVYIGTRMSPPA		1042
Db	1935	PLPSKFA-----SRPASVAESIKDAEAKSKES-----RRESVAEKS		1971
QY	1043	PPPOQRVEQGHGKTRATGHYNTISRMDDRRVMDDHYSDBDRDCEADROPYHRSRSTEO		1102
Db	1972	PLPSK-----EASRPASVAESI-----KDEAKSKESRRESVAEK		2007
QY	1103	RPLLETTTTRSRS-----SERPDNLMRSMPSLIMGRSAPSPALSRSHPTGTSVOTSES		1157
Db	2008	SPLPSKEASRPASVAESIKDAEAKSKESRRESVAEKSPPLPSKEASRPASVAESIKDEAE		2067
QY	1158	STGTGTRGKQLPOLPPKGTLESADIDIERNRQMKLNTKYQVAGSDPRLQDYHSHKYSR		1217
Db	2068	KSKSEESR-----ESAEEKSPLPSKASRPASVAESVKDEADKSKESRRESVAES		2118
QY	1218	GMPHRRGADTVSTKSDSDVDVSAVSRITSSASRFSSTYSMVQSERPRGNKRTISVFTSK		1277
Db	2119	G-----KMQSIKGDGSPLEKVSRESPEVAE-----SVKDDPVK-----SK		2152
QY	1278	MONROMGVSGKNLTKTSTISGDMCSLEKNGSODTAVGALGTSGKKRRSIGAKVAIV		1337
Db	2153	EPSRRESVAES--VTAADARDSDPLSKSGKASRESVYDVKDEAEKQESRREKRTISVI		2210
QY	1338	GLSKR-SRSASQLSQTQGGGKKLRSIVQSTETGLAVEMRNMMTRQASRBSSTDSNNSTYS		1396
Db	2211	PPKAKDKSPREVLQVPSMTETIREDADQPKPQQA-----ESRRESIAESIKAASS		2261
RESULT 11				
T42727				
proliferation potential-related protein - mouse				
C:Species: Mus musculus (house mouse)				
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000				
C:Accession: T42727				
R:Witte, M.M.; Scott, R.E.				
submitted to the EMBL data Library, November 1998				
A:Reference number: 222246				
A:Accession: T42727				
A:Status: preliminary: translated from GB/EMBL/DBJ				
A:Molecule type: mRNA				
A:Residues: 1-1560 <MIT>				

A;Cross-references: EMBL:U08913; NID:9385884; PID:9385885; PIDN:AACT2432.
A;Experimental source: Strain Balb/C
C;Genetics:
A;Gene: P2P-R
C;Function:
A;Description: involved in hnRNP association and Rb1 binding
C;Superfamily: RING finger homology
E;57-107/Domain: RING finger homology <RRR>

Query Match	3.2%	Score 266	DB 2:	Length 1560
Best Local Similarity	19.7%	Pred. No. 1.9e-05		
Matches 257	Conservative 176	Mismatches 447	Indels 422	Gaps 61
Qy	196	NTTLQPOQKPVYRGRLNEAPAEQEKAKLHEQOFQGFAGFDLSPVAPVEKRAHGLTRQDITI	255	
Db	427	NTM-PTTQAPLRSREFFYRQKLLKEESKPPYGGSSYSRSYIDSKRS-GSTR----	479	
Qy	256	KNGSGVKHQIADSMEDKRRSPSVSRDQNRREYGE-----EREDYQYVPSG----	304	
Db	480	-----SRYSKRSFSRSHRSYRSRPPYRGRGRKSRNRSRSHGYHR	523	
Qy	305	TMPKPSDYADRRSQREQFY--EERGHLYN-RDSNRGHRISKEYIVDEDEVS-----	356	
Db	524	SRSRPPPRRYHRSRSDQAFRGQSPTRKNNPGRGETEREYFNRYREVPPTYDIKAYGHS	583	
Qy	357	---RDEYRORREERYQARYS--DPLNARYVKKPOPYEQGR-----IHAE	398	
Db	564	VDFRDPFEKE-RYRWEKRYKRYTKYKGYAVGQAPRPSANREDESPRLPLNIRNSP	642	
Qy	399	VSRAHE-----RRSDVSLAN-AELSDSRISLIMDRPSRORSYSE-----	439	
Db	643	FTRGREDYVAAQSHRNNTLGGVNEYEKLISTRSHNAKDPAKKESESEVPGDGKGNKHK	702	
Qy	440	---RAAMENORYS-----METTREAOGSS-----YPORTSNHSPTP-R	477	
Db	703	KHRKRNEKEGESESLNPELLGFRKCRQSGSIDETKDTPLFVFPSPR---EDATPYR	758	
Qy	478	RSPILDRPDMRRADSLRKQHHLDPSAAVRKTKRE-----KMETMLR-----NDLSLS	525	
Db	759	DEPMALSTTKSVSDKDKREKDKPKVYSDTKTKRKSQSGATKKKDVNLKPSKGPQEKVVG	818	
Qy	526	DQSESVRRPPPRPHSKKGGKMOVSLSSS---EEELASTPEYTSODVLELSES-----	577	
Db	819	DREKPSRSPPLKAKKAEATKIDYKVPSSSQKDEKVTGTTPRAHSAKSAKEHQEAKPAD	878	
Qy	578	-----VSEGDQSGKGRKTKSEQCVLSDSNTRSRQKKRMATYGGHLSLEEDLWSE	626	
Db	879	EKKVKDCSDKIDISEKPPASDEKAKKPEKNKLLDS--KGKRRKRT--EKKSVYDKPESSS	934	
Qy	627	POI-KDSQVD-----TCSST-----LNEEH-----SHD	650	
Db	935	MKISVGEGLVPSPPKRMMEGDEKLEFTRPKDKTASSTTPAKKTKLNRFGKTKIGNE	994	
Qy	651	KHPVWQPSKDDRLIGRILLKRLKDSVPPRDSGAMLLKVGKMTPESSGRLCATITVY	710	
Db	995	NASTKPESEKLETSRSKTKQEK--VKGAAR-----KXAGSGSSGSS-----TLV	1037	
Qy	711	KKGLADTVGH-LRPGDEVLEWNGRLLQATGEVYNNILLESKPEFOVELVYSRPIGDI	769	
Db	1038	DYTSSTSTGSGSVKRSSEKTDTRTVIK--TMEEVYNN--NTAPADVILIMIHV-----	1088	
Qy	770	RIPDSTHOLESSSSSEFEQKMD---RPSIVTSPMS-----	803	
Db	1089	-----QSKMDKDFSEEBEDVKTTPQIOSGKPSIILKNVTKPSATAKYTEKESQ	1140	
Qy	804	PGLMDVPOPLSGOLSIKLMFDKVGHQLLVTLIGAKDLPSREDGRP-----NP-	852	
Db	1141	PKLQKLPKEASHL-----MQHEL-----RSSKGSASSSEKGRAKKDRHSSSEKMDND	1188	
Qy	853	YKTIYFLPRS-----DKKRRKTKYVK--TEPKNNQFTI-----	886	
Db	1189	KRKSQAOPKSESTVDRLSEQGHFKTLQSGSKETPRSEKHEHVSAGSSNKQFTTPGRDKVDY	1248	

```

OY 887 ---YSPVNRREFR-----EMLTTLMDQARVREES--EFLGELLLELALL 930
Db 1249 DSRDYSSSKRRBERGELARRKDSPPRGKESLSGQSKSLREERLPPKGAASSKKNSSPPR 1308
OY 931 DDEPHWYKQJTDHVSLSPLRPBSPYLRPRROHGESPRRL---QRKSRIIDSEVSDYCE 987
Db 1309 DKPRPDHK-----APYETKRCEETKRYDKLSGKERENHADA----- 1346
OY 988 DGVGVVSDYRHNGRDLQSSTLSVPEQVMSSNIC--SPSGSPHNVY-IGTRSMSPSAPP 1044
Db 1347 -----RNGKD--SSGKPLP-----CILNPDRLMEKELAVGQVEK-SAVKPK 1385
OY 1045 PQRNVBOGHRGTRATGHYUTISMDNHRVWDHNTSSRDQDCSAADQRPYRNRSTE--- 1101
Db 1386 PQ-----LSHSR-----LSSDLTETLEAAREPYNESDSNNV 1420
OY 1102 -----ORPLERTTTRRS--ERP-DTNLMSMPMLTGRSAPRPPAL 1142
Db 1421 SVKEEAVASISKOLKEKTEKKESLITVAQAQFCDABRQSQSPSVSRKSNISPGSQ 1480
OY 1143 SKSHPRTSVQVSPSSPTPGTRGRGRLPOLPRPKGLERSAMDIERNQKMLKKYQVAG 1202
Db 1481 TSHSSASSASAGQDSKK-----KKKKKKKKKKKKKKKKKKHAG 1522
OY 1203 SPPRLQDYHNSKTRSGMDHNRGADLYTSSKSDSDVSYSVAVS 1244
Db 1523 AGDVYKESQKHKK-----KKKKKKKKDKKKEDDKVRNST 1559

```

RESULT 12
 T20531
 hypothetical protein F07A11.6a - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T20531, T27776
 R:Palmer, S.
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z19287
 A:Accession: T20531
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2526 <W12>
 A:Cross-references: EMBL:Z66511; PIDN:CAB54210.1; GSPDB:GN00020; CESP:F07A11.6a
 A:Experimental source: clone F07A11
 R:Gajadsky, S.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z20417
 A:Accession: T27776
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2526 <W12>
 A:Cross-references: EMBL:Z69904; PIDN:CAB54501.1; GSPDB:GN00020; CESP:F07A11.6a
 A:Experimental source: clone ZK20
 C:Genetics:
 A:Gene: CESP:F07A11.6a
 A:Map position: 2
 A:Introns: 36/2; 92/2; 182/2; 272/3; 344/2; 426/2; 490/1; 541/2; 700/3; 770/3; 1286/3

Query Match	3.2%;	Score 266;	DB 2;	Length 2526;
Best Local Similarity	17.9%;	Pred. No. 3.5e-05;		
Matches 298;	Conservative 235;	Mismatches 592;	Indels 540;	Gaps 66;
QY	180	QGEILLTSGAMWYNNSSGNTLQDPDQKVPRLGLNEAPQCKKKKLLHQPQFGAGPGLSVLP	239	
	:	:	:	:
DB	908	RHDITLTKSGP---SNAITNLQARSGSMTSGDEKKSAP-----STPVVRAGSDLVAQ	956	
QY	240	AVEKGRGHILTRQDTITKNSGVKHOIASDMP--SDRRKSPSY-----SRPDN	284	
	:	:	:	:
DB	957	IMSNDPGLGRLPLPREKKSSALQNTQNHQPHSNANSTPSTPSTQIQAAMKKDKEKEKK	1016	
QY	285	RYEODESEEREDYSQYVPSDGTMPRSPSDVADRSSQHEPQFYEPGHLNADSNRRQHRHS	344	


```

Db 194 PA-----RAPARGMEDRRPPGKPGDILTAPGNG 224
QY 332 NY-----RDSNRGRHRSKEYIYDDEDEVEDERQ---RREEYQA 371
Db 225 SHGPPTRRASEARMASTARDEGMDHAG-----GGTGDTSRSPAGLRANSYQA 274
QY 372 RYSDPMLARYPVKPOP-----YEQMRIHAEVSRAHRRHSDVSLA 414
Db 275 A-RPARAPVSPAPQVPQCPGPGSRATPGCFRPEQST-----EAPSPDPGYP 323
QY 415 NAELEDSRISLLRMDRPSKORSVSERRAMENORSYSMERTREAGOGSSYPORTSNHSP 474
Db 324 GAV-----APARE-----ERTGPAGGFQAAPHTAPY5QA 353
QY 475 TPRRSPITLDRPDMRRADSLRKOHILDPSSAVKRTKREKMETMLRNDLSSD-QSESVAP 533
Db 354 APARQPPAREEE-----EANSYSDQA--TTIGALEFSLIYDQNSNLOCTITRA 403
QY 534 PPPRPHSKKGGKMRQVLSSESEELASTPEYTCDDVELESESVSEKGDOSQKRRKTSE 593
Db 404 KGLKPMDSN-----GLADPYVKLHLPP-----GASKSNKLRTK- 436
QY 594 OGVLSDNTRSEROKRMYYGGHSLLEDLEWSEPOIKDSGVDTCSSTTLNEBHSXKHP 653
Db 437 -----TLNTRNPNVNETLQYHGIT-EEDM-----QRKTLRISVCD---EDKFGHNE-- 479
QY 654 VTWQPSKDGRLIGRILNLRKLDGSVPRDSGAMGLKVGCGKMGESGLCAFITKVKKG 713
Db 480 -----FTGE-----TRPSLKLKLANOKRKNNTIC--LERVITPM 509
QY 714 SLADTVGHLRPGDEVLWNGRLLOGATFEYVYNIILESKRPEOVELVVSRIPIIDIP 773
Db 510 KRAGTGSAR-----GMALYE-----EEQVE-----RIGDIEE--- 537
QY 774 SYHAQLESSSSSFESQKMDRPSISVTSPMSPGMLRDVPOFLSGQLSIKLMDKVGHQLIV 833
Db 538 -----RGLITVLSMYSTQOGGLIYI 556
QY 834 TILGAKLPRREDGRPNPYKIYFLPDRSDKNKRRKTQVTKLEPKNNOPTIYSPVHR 893
Db 557 GILRCVHLAAMDANGSDPFVKMLKPDMGKRAKHKTKQIKKTLNPEFNEFFYDIKH-S 615
QY 894 EFERMLEITLMDQARVREESEFLGELLILETALTALDDEPHWK 938
Db 616 DLAKSLDISVMDYD--IGKSNDYIGCQQLGI-SAKGERLKHMYE 657

RESULT 14
158166
rabphilin-3A - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
C:Accession: I58166
R:Li, C.; Takei, K.; Geppert, M.; Daniell, L.; Stenius, K.; Chapman, E.R.; Jahn, R.; De
Neuron 13, 885-898, 1994
A:Title: Synaptic targeting of rabphilin-3A, a synaptic vesicle Ca2+/phospholipid-binding
A:Reference number: I58166; MUID:95033210
A:Accession: I58166
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-684 <RES>
A:Cross-references: EMBL:U12571; NID:9533710; PIDN:AAA62662.1; PID:9533711
F:376-492/Domain: protein kinase C C2 region homology <KC2A>
F:534-649/Domain: protein kinase C C2 region homology <KC2B>

```

Query Match 3.2% Score 262.5; DB 2: Length 684;

Best Local Similarity 18.8%; Pred. No. 9.9e-06;

Matches 177; Conservative 105; Mismatches 236; Indels 363; Gaps 38;

QY 42 AWMORKEEKEEOSVLKKEEHKAQPTWPFEGSITELVNVNLOPOOK-QPNEKEPQTK 100

```

Db 35 AOTDRORKEE-----LTDEEK-----EILNRVIARAEMKMETQEORIGR 74
QY 101 LHQOFEMYKQYVKKMGE-----SQQQOEOQKDAPTGICIKHTRKADCGHNCSCQTKFC 156
Db 75 LVDRLETKRNKNVAGDGYNRCITLCEBQGLMSA--CVVCEQD-----CKKNVC 119
QY 157 ARCGHVSLSRNSNVMVWCNLCRKQOEILITKSGAMFYNSGNTLQOPQKVRGRLNEAP 216
Db 120 TKCGVETSNRRPRPVMLCKLCLEOREYMKRSGAMF-----KGFPPKQVLP 164
QY 217 QEKKAALHEOPQFOGABGLSVPAVEKRAHGLTRQDTIKNGSGVKKHQLASMDPSDKRS 276
Db 165 QPMPIKTKPQOPAGEBA-----TQEQPTPESRH 193
QY 277 PSYSRQONRRYEDSEEDYQYVPSDGTMPRSR--DYADRSQRE--PQYFEPQHL 331
Db 194 PA-----RAPARGMEDRRAPGKPGDILTAPGNG 224
QY 332 NY-----RDSNRGRHRSKEYIYDDEDEVEDERQ---RREEYQAR 373
Db 225 SHGPPTRRASEARMASTTRDSEGDHGHG-----GGAGDTSRSPGEGGLRANSYQAS- 278
QY 374 RSDPMLARYPVKPOP-----YEQMRIHAEVSRAHRRHSDVSLANA 416
Db 279 RPARASMPSPAPQVPQCPGPGSRAAPGCFRPEQST-----EAPSPDPGYPGA 328
QY 417 ELEDSTRISLLRMDRPSKORSVSERRAMENORSYSMERTREAGOGSSYPORTSNHSPPT 476
Db 329 V-----APARE-----ERTGPTGFGQAAPHTAPY5QAAP 358
QY 477 RSRPIPLDRPDMRRADSLRKOHILDPSSAVKRTKREKMETMLRNDLSSD-QSESVRPP 535
Db 359 ARQPPAREEE-----EANSYSDQA--TTIGALEFSLIYDQNSNLOCTITRAKG 408
QY 536 PRPHSKKGGKMRQVLSSESEELASTPEYTCDDVELESESVSEKGDOSQKRRKTSEOG 595
Db 409 LKPMDSN-----GLADPYVKLHLPP-----GASKSNKLRTK- 439
QY 596 VLSDNTRSEROKRMYYGGHSLLEDLEWSEPOIKDSGVDTCSSTTLNEBHSXKHPVT 655
Db 440 --TLNTRNPNVNETLQYHGIT-EEDM-----QRKTLRISVCD---EDKFGHNE--- 482
QY 656 WQPSKDDRLIGRILNLRKLDGSVPRDSGAMGLKVGCGKMGESGLCAFITKVKKSL 715
Db 483 -----FTGE-----TRPSLKLKLANOKRKNNTIC--LERVITPKR 514
QY 716 ADTVGHLRPGDEVLWNGRLLOGATFEYVYNIILESKRPEOVELVVSRIPIIDIP 775
Db 515 AGTGSAR-----GMALYE-----EEQVE-----RIGDIEE--- 540
QY 776 HAQLESSSSSFESQKMDRPSISVTSPMSPGMLRDVPOFLSGQLSIKLMDKVGHQLIYTI 835
Db 541 -----RGLITVLSMYSTQOGGLIYI 561
QY 836 LGAKDLPREDGRPNPYKIYFLPDRSDKNKRRKTQVTKLEPKNNOPTIYSPVHRREF 895
Db 562 IRCVHLAAMDANGSDPFVKMLKPDMGKRAKHKTKQIKKTLNPEFNEFFYDIKH-SDL 620
QY 896 RERMLEITLMDQARVREESEFLGELLILETALTALDDEPHWK 938
Db 621 AKKSLDISVMDYD--IGKSNDYIGCQQLGI-SAKGERLKHMYE 660

```

RESULT 15

A48097

rabphilin-3A - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A48097

R:Shitataka, H.; Kaibuchi, K.; Sakoda, T.; Kishida, S.; Yamaguchi, T.; Wada, K.; Miya

Mol. Cell. Biol. 13, 2061-2068, 1993

A:Title: Rabphilin-3A, a putative target protein for smg p25A/rab3A p25 small GTP-bln

A:Reference number: A48097; MUID:93204952

A:Accession: A48097
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-704 <SH1>
A:Cross-references: GB:D13613; NID:g285645; PIDN:BA02780.1; PID:d1003285; PID:g285646
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:127629)
C:Superfamily: protein kinase C C2 region homology
F:396-512/Domain: protein kinase C C2 region homology <KCA>
F:554-669/Domain: protein kinase C C2 region homology <KC2B>

Query Match 3.28; Score 261.5; DB 2; Length 704;
Best Local Similarity 18.38; Pred. No. 1.2e-05;
Matches 172; Conservative 112; Mismatches 307; Indels 351; Gaps 38;

```
QY 45 DRQKEEKEQSVLKIKKEHKAQPTQWPFSGITELVNNVLOPQOKPNEKEPQTKLHQ 104
DB 42 DRQKQKE-----LTDEEK-----EINRVIAAEKME-----69
QY 105 FENYKEQVKKMGESQOQEO-KGDAPL-CGICHT--KPADCGHNCYCOTKFCARCG 160
DB 70 -EMQERIGRLVDRLNNKKVAGDVNRCILGEOIGMSAC--VVEDCKKNVCTKCG 127
QY 161 GRVSLRBNKVMWVNCICRQOEILTKSGAMFVNSGNTLOPDQKVPKGLRNEEAPQEK 220
DB 128 VETSNRRPHFWLCKICEQREVWKRSGAMFEKG-----FPKQVLPQ-----PMPIK 174
QY 221 AKLHEQPOFQAGPDLVAVAY--EKRAHGLTQDTIKNGSGVKHQIASDMPDRKR-- 275
DB 175 KNRQOQVSEVPA-APRPAPEPKHAPAPTRGDT-EDRRGPGQKTGDMTSAPGRGSY 232
QY 276 SPVSVRQNNRYEQQEEDYQYVPSDGTMPRSPDIADRRSQREPOFYEEPHLNYRD 335
DB 233 GPVVRRASEAR--WSSSGRSDSDMDQGHMAAGDS-----QSPAGL-----272
QY 336 SNRGHRHSKXYIYDDEVESNDEYERQREERYQARYRSDPNLARYPYKQPYEQMRT 395
DB 273 -----RANSVQAS--RPAPASMGSPAPQP-----296
QY 396 HAEVSRARHRRHSDVSLANAELEDSTRISLMDRRPSRQSVERRAMENORSYMER 455
DB 297 -----296
QY 456 REAGQSSYPORTSNHSPTPRSPPIPLDRPMRADSLRKQHLDSSAVRKTREKME 515
DB 297 ---GQPPPGSGSRSPGPTGR--FPQRPVAPSDP-----DYGAAPRREERTG 342
QY 516 TMLRNDLSLSDQSESVRPP-----PPRPHKSKGKGKMOVLSLSS 555
DB 343 GIGGYSAAGTREDRAGHPRPGSYTOASAAAPQPVASARQPPPEDELEAN---SYDSD 398
QY 556 EEELESTPEY--SCDVELESESVSEK---DSQ-----KGRKTSQGV 596
DB 399 EATTLGALEFSLYDQDNSSLCTTIKAKGLKPMDSNGLADPYVKLHLPLGASKSNKLR 458
QY 597 LSDNTRESEKQKKRMYGSHSLLEDLEWSEPOIKDSCVDTCSSTILNEHSHDKHPVTW 656
DB 459 KTLRNTNPNINMETLVYHGIT--DEDM---ORKTLRISVCD---EDKFGHNE-----502
QY 657 QPSKDDRLIGRLINKRLKDGSVPRDSGAMGLKVVGGKMTESGRICAFITVKKGSILA 716
DB 503 -----FIGE-----TRFSIKKLKPNQKRNFNIC--LERVIPMKRA 535
QY 717 DTVGHLRPGDEVLEWNGRLLOGATEEYVNIILESKPEPOVELVSRPIGDIRIPDSTH 776
DB 536 GTTGSAR-----GMAIYE-----EEOVE-----RIGDIE-----560
QY 777 AOLSSSSSFESQKMDRPSISTSPMSPGMLRDVPOFLSGOLSIKLMFDKVGHLIVTIL 836
DB 561 -----RKTLVSLMTSTOOGGLIVGIT 582
QY 837 GAKDPSREDGRPNRPVYKIFLPDRSDKNKRTKTVKTLPEPKMNQTFIYSPVHRRER 896
```

```
DB 583 RCYHLAMANGYSDFEVKLMKPRDMGKRAKHKTQIKKTLNPEFNEEFYDIKH-SDLA 641
QY 897 ERMLETTMDQAVRERESEFLGEILILETALIDDERPHWK 938
DB 642 KKSIDISVWDYD--IGKSNIDYIGGCOQIGI-SAKGERLKHWE 680
```

Search completed: November 21, 2001, 16:06:36
Job time: 182 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:00:09 ; Search time 29.05 seconds
(without alignments)
3318.146 Million cell updates/sec

Title: US-09-617-099b-1
Perfect score: 8285
Sequence: 1 MSAPGPRGRPAAPTAAASQP.....TRPASQSLSTGTPSYRS 1590

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_0601.*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
23: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1586.5	19.1	374	19	AAW29640
2	1015.5	12.3	237	21	AAB34848
3	1010.5	12.2	237	21	AAB34847
4	624	7.5	134	21	AAB34849
5	566	6.8	128	21	AAB34850
6	420	5.1	126	21	AAB40725
7	417	5.0	86	22	AAB55848
8	417	5.0	86	22	AAB57639
9	417	5.0	86	22	AAB58049
10	307	3.7	1898	20	AAV30795
11	270.5	3.3	3266	21	AAB42491

12	258.5	3.1	2742	21	AAB23012
13	258.5	3.1	2842	15	AAAR6308
14	258.5	3.1	2843	16	AAW11922
15	258.5	3.1	2843	19	AAW76140
16	258.5	3.1	2843	19	AAW76144
17	258.5	3.1	2843	21	AAAB23011
18	258.5	3.1	2973	19	AAW76821
19	258.5	3.1	2973	21	AAW70304
20	258.5	3.1	2973	22	AAW72782
21	256.5	3.1	704	15	AAAR57421
22	253.5	3.1	1341	15	AAAR5657
23	253.5	3.1	2843	15	AAAR5634
24	250.5	3.0	2843	13	AAAR6052
25	250.5	3.0	2843	18	AAW35392
26	250.5	3.0	2843	19	AAW38370
27	250.5	3.0	2860	15	AAAR6307
28	248.5	3.0	2343	21	AAAB12453
29	243	2.9	1596	18	AAAR61347
30	235	2.8	2432	21	AAV55555
31	230.5	2.8	2453	21	AAAB12454
32	230	2.8	2819	22	AAAB35408
33	226	2.7	2101	21	AAV49936
34	225.5	2.7	1780	19	AAW35863
35	225.5	2.7	1780	21	AAAB15380
36	222	2.7	2101	15	AAAR47173
37	220	2.7	2115	21	AAV49937
38	217	2.6	2192	21	AAV53920
39	217	2.6	2192	18	AAW21732
40	217	2.6	2272	22	AAW21731
41	212.5	2.6	1054	12	AAAB6897
42	210	2.5	2213	20	AAV6079
43	209	2.5	1239	20	AAV55931
44	208	2.5	2442	21	AAV77575
45	207.5	2.5	1064	22	AAAB6898

ALIGNMENTS

RESULT 1	
AAW29640	AAW29640 standard; Protein: 374 AA.
XX	
AC	AAW29640:
XX	
DT	09-NOV-1998 (first entry)
XX	
DE	Human secreted protein C0618_1.
XX	
KW	Secreted protein: C0618_1; human.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 100 /note= "encoded by WAA"
FT	Misc-difference 263 /note= "encoded by TGS"
FT	Misc-difference 289 /note= "encoded by YTA"
FT	Misc-difference 372 /note= "encoded by TYT"
FT	Domain 350 /note= "predicted transmembrane domain is centered around amino acid 350"
XX	
PN	WO9831802-A1.
XX	
PD	23-JUL-1998.
XX	
PF	21-JAN-1998; 98WO-US01007.
XX	
PR	20-JAN-1998; 98US-0010047.

21-JAN-1997; 97US-0072103.
16-JUN-1997; 97US-0877035.
(GEMV) GENETICS INST INC.
Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
Racle LA, Spaulding V, Treacy M;
WPI: 1998-414101/35.
N-PSDB; AAV40485.
New isolated polynucleotides and secreted proteins - obtained from
human foetal kidney, adult testes, placenta, adult brain and foetal
brain cDNA libraries
Claim 22; Page 71-72; 104pp; English.

Query Match	19.1%	Score 1586.5;	DB 19;	Length 374;
Best Local Similarity	85.8%	Pred. No. 2e-103;		
Matches 321; Conservative	6;	Mismatches 6;	Indels 41;	Gaps 1;
QY 1258	MSVQSERPGRNRKISVFNFSKMNROMGVSGKNILTKSTISIGDMCSLEKNDGSQDPAVGA	1317		
Db 1	msvqserpggnkktisvftskmgstqmgisgkmktstisgdmcslekdngdsqstavyt	60		
QY 1318	LGISGKKRRRSITGAKMAVAIVGLSRKSRASASOLQTEGGGKKLRSTVQBSTETGLAVEMRN	1377		
Db 61	lgtsgkkrtrssalgaamvaivglstrksrassqstleaagxkrlrstvtqrstetglavemrn	120		
QY 1378	WMTRQASRESRNDGSNNVSSSGNLIFFPVRRLASDSQFEDLDGCPADLVNGQTLATPAM	1437		
Db 121	wmttrqastescdgsmnsyssegnliffpvrlasdsqfiedglgpaqlvgrqlatpam	180		
QY 1438	GDIOVGMDKKGOLEVEIIRARGLVKSGSKTLPAPYKVVLLDNGVC-----	1485		
Db 181	gdiovgmddkkqgleveirarglvvkpsktlpapykvvllndvgtlakkqglevei	240		
QY 1486	-----IAKKKTIVARKTLEPLLYOQLLFEEPSDGRV	1516		
Db 241	rarglvvkpsaktlpapykvvllndvgtlakkktkvarktlrplyqllsfespgkv	300		
QY 1517	LQITVWDGYGMDHKSFFMGVAQOILLDELELSNMVIGWKLEPPSSLVPTAPLTRASQ	1576		
Db 301	lqitlvwdygmhksffmgvaqilldelelsnmvlgwkllppsslvptapltrrsq	360		
QY 1577	SSLESTGSPYSRS 1590			
Db 361	ssleestgspysrs 374			

ID	AAB34848
AC	AAB34848 standard; Protein; 237 AA.
XX	
XX	AAB34848;
DT	26-JAN-2001 (first entry)
XX	
DE	Human secreted protein sequence encoded by gene 46 SEQ ID NO:136.
XX	
KW	Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW	nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KM	antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;
KW	antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KM	cancer; immune disorder; cardiovascular disorder; wound healing;
KN	neurological disease; infectious disease; chromosome identification;
KW	Chromosome 6.
XX	
OS	Homo sapiens.
PN	WO200058356-A1.
PD	05-OCT-2000.
XX	
PF	22-MAR-2000; 2000MO-US07535.
XX	
PR	26-MAR-1999; 99US-0126511.
PR	17-DEC-1999; 99US-0172413.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Ruben SM, Komatsoulis G;
XX	
DR	WPI: 2000-594639/56.
XX	
PT	Fifty nucleic acid molecules encoding human secreted proteins, useful
PT	in the prevention, treatment and diagnosis of cancer, immune disorders,
PT	cardiovascular disorders and neurological diseases -
PT	
PS	Disclosure; Page 419-420; 425pp; English.
XX	
CC	The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC	human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
CC	AAB34822 represent human secreted polypeptide sequences and proteins
CC	homologous to them, which are given in the exemplification of the present
CC	invention. Human secreted proteins have activities based on the tissues
CC	and cells the genes are expressed in. Examples of activities include:
CC	cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
CC	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
CC	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC	cardiant. The polynucleotides and polypeptides are useful for
CC	preventing, treating or ameliorating a medical condition in e.g. humans,
CC	mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC	polypeptides can also be used as a food additive or preservative to
CC	increase or decrease storage capabilities. The polynucleotides are
CC	useful for chromosome identification. They are also useful as probes for
CC	diagnosing a disorder related to the female reproductive system,
CC	particularly breast and/or ovary cancer. They are also useful in the gene
CC	therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC	agonists and antagonists from the present invention are useful in the
CC	diagnosis, treatment and prevention of cancer, immune disorders,
CC	cardiovascular disorders, wound healing, neurological diseases and
CC	infectious disease. AAC59967 to AAC59965 and AAB34772 represents sequence
CC	used in the exemplification of the present invention.
XX	
XX	
Sequence	237 AA;
30	

	Query Match	12.3%	Score 1015.5;	DB 21;	Length 237;
	Best Local Similarity	82.8%;	Pred. No. 1.4e-63;		
	Matches 197;	Conservative 19;	Mismatches 21;	Indels 1;	Gaps 1
QY	1353	EGGCKKLASTYVORSTETIGLAVEMRMWMTFQASRESTDGSMNYSSEGCNLIFFGVRLASDS	1412		

Db 1 esghkklkstlqstetgmaemrk-mvrgpserstdgsinsyssegnlfpvgvlgads 59
 QY 1413 QESDFLDGLGPAOLVGRDTLATPAMGDIQVGMMDKRGQLEVEIIRARGLVVRGSKTLP 1472
 Db 60 qfsdflgdlgpaqlvgrqlatpamgdlqgmedkkgglevevitarstlqkpskstpa 119
 QY 1473 PYKVYVLLDNCVCIAKKRTKARKTLEPLVQQLSFESPOGRCVLIQIYWGDIYGRMDHKS 1532
 Db 120 pykvvylldngaciacakkttriarctldpdyqslvfdespgkqvivwgdymdhkc 179
 QY 1533 FMGVAQIILDELEISNMVIGWFKLEPPSLVDPSTAPLTRRASQSSLESSTGPGYSRS 1590
 Db 180 fmgvaqllleeldlssmviywkklfpsslvdpdltrrasqsslesstgpcirs 237

RESULT 3
 ID AAB34847 standard; Protein: 237 AA.
 AAB34847;
 DT 26-JAN-2001 (first entry)

Gene 46 human secreted protein homologous amino acid sequence #135.
 Human: secreted protein; diagnosis; cytostatic; immunosuppressive;
 nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 antidiabetic; antinflammatory; antitumor; vulnerable; anticonvulsant;
 antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
 cancer; immune disorder; cardiovascular disorder; wound healing;
 neurological disease; infectious disease; chromosome identification.
 Rattus norvegicus.
 OS
 XN
 XX W0200058356-A1.
 PD 05-OCT-2000.
 PF 22-MAR-2000; 2000WO-US07535.
 PR 26-MAR-1999; 99US-0126511.
 PR 17-DEC-1999; 99US-0172413.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM, Komatsoulis G;
 DR WPI; 2000-594639/56.
 XX
 DR FIFTY nucleic acid molecules encoding human secreted proteins, useful
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases -
 PS
 XX Disclosure: Page 418-419; 425pp; English.

The polynucleotide sequences given in AAC59966 to AAC60015 encode the
 CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
 CC AAB34852 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities include:
 CC cytostatic; immunosuppressive; nontropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
 CC vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for
 CC preventing, treating or ameliorating a medical condition in e.g. humans,
 CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities. The polynucleotides are
 CC useful for chromosome identification. They are also useful as probes for
 CC diagnosing a disorder related to the female reproductive system,
 CC particularly breast and/or ovary cancer. They are also useful in the gene
 CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,

CC agonists and antagonists from the present invention are useful in the
 CC diagnosis, treatment and prevention of cancer, immune disorders,
 CC cardiovascular disorders, wound healing, neurological diseases and
 CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 237 AA;

Query Match 12.2%; Score 1010.5; DB 21; Length 237;
 Best Local Similarity 82.4%; Pred. No. 3.2e-63;
 Matches 196; Conservative 20; Mismatches 21; Indels 1; Gaps 1;

QY 1353 EGGGKILRSTVORSENTGLAVERMWMTRQASRESTDGSMNSSEGNLFPVGRASDS 1412
 Db 1 esghkklkstlqstetgmaemrk-mvrgpserstdgsinsyssegnlfpvgvlgads 59
 QY 1413 QESDFLDGLGPAOLVGRDTLATPAMGDIQVGMMDKRGQLEVEIIRARGLVVRGSKTLP 1472
 Db 60 qfsdflgdlgpaqlvgrqlatpamgdlqgmedkkgglevevitarstlqkpskstpa 119
 QY 1473 PYKVYVLLDNCVCIAKKRTKARKTLEPLVQQLSFESPOGRCVLIQIYWGDIYGRMDHKS 1532
 Db 120 pykvvylldngaciacakkttriarctldpdyqslvfdespgkqvivwgdymdhkc 179
 QY 1533 FMGVAQIILDELEISNMVIGWFKLEPPSLVDPSTAPLTRRASQSSLESSTGPGYSRS 1590
 Db 180 fmgvaqllleeldlssmviywkklfpsslvdpdltrrasqsslesstgpcirs 237

RESULT 4
 ID AAB34849 standard; Protein: 134 AA.
 AAB34849;
 DT 26-JAN-2001 (first entry)

Gene 46 human secreted protein homologous amino acid sequence #137.
 Human: secreted protein; diagnosis; cytostatic; immunosuppressive;
 nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 antidiabetic; antinflammatory; antitumor; vulnerable; anticonvulsant;
 antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
 cancer; immune disorder; cardiovascular disorder; wound healing;
 neurological disease; infectious disease; chromosome identification.
 Rattus norvegicus.
 OS
 XN
 XX W0200058356-A1.
 PD 05-OCT-2000.
 PF 22-MAR-2000; 2000WO-US07535.
 PR 26-MAR-1999; 99US-0126511.
 PR 17-DEC-1999; 99US-0172413.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM, Komatsoulis G;
 DR WPI; 2000-594639/56.
 XX
 DR FIFTY nucleic acid molecules encoding human secreted proteins, useful
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases -
 PS
 XX Disclosure: Page 420-421; 425pp; English.

The polynucleotide sequences given in AAC59966 to AAC60015 encode the
 CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
 CC AAB34852 represent human secreted polypeptide sequences and proteins

CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neutrotropic; neuroprotective; antiviral;
CC antidiabetic; hepatotropic; antidiabetic; antiinflammatory; antilucer;
CC vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.

XX Sequence 134 AA;

Query Match 7.5%; Score 624; DB 21; Length 134;
Best Local Similarity 85.1%; Pred. No. 2e-36; Indels 0; Gaps 0;

Matches 114; Conservative 11; Mismatches 9;

QY 825 DKVHOLIVTLGAKDLPREDGRPNPVYKTYLPDRSDKNKRRTKYKTKLEPKWNOT 884

DB 1 dkvghqlivmqlgtldprvdgrpnpykmyflpdrskskrtktvkllepwmqt 60

QY 885 FYSPVHRREFRERMLFTLMDQARVRESEFGLGELLLELTALDDEPHWYKLOTNDV 944

DB 61 fygshvhrdrtremleltvwdqprvgeeseflgellleltallddephwylqthde 120

QY 945 SSLPRRSPYLP 958

DB 121 sslprgppsimpr 134

RESULT 5
AAB34850 standard; Protein: 128 AA.

XX AAB34850;

DT 26-JAN-2001 (first entry)

DE Human secreted protein sequence encoded by gene 46 SEQ ID NO:138.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX neutrotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antiinflammatory; antilucer; vulnerary; anticonvulsant;
XX antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
XX cancer; immune disorder; cardiovascular disorder; wound healing;
XX neurological disease; infectious disease; chromosome identification;
XX chromosome 6.

XX Homo sapiens.

PN W0200058356-A1.

PD 05-OCT-2000.

PF 22-MAR-2000; 2000WO-US07535.

PR 26-MAR-1999; 99US-0126511.

PR 17-DEC-1999; 99US-0172413.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI, 2000-594639/56.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful
XX in the prevention, treatment and diagnosis of cancer, immune disorders,
XX cardiovascular disorders and neurological diseases -

PS Disclosure; Page 421; 425pp; English.

XX The polynucleotide sequences given in AAC59966 to AAC60015 encode the
XX human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
XX AAB34852 represent human secreted polypeptide sequences and proteins
XX homologous to them, which are given in the exemplification of the present
XX invention. Human secreted proteins have activities based on the tissues
XX and cells the genes are expressed in. Examples of activities include:
XX cytostatic; immunosuppressive; neutrotropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antilucer;
XX vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
XX cardiant. The polynucleotides and polypeptides are useful for
XX preventing, treating or ameliorating a medical condition in e.g. humans,
XX mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities. The polynucleotides are
XX useful for chromosome identification. They are also useful as probes for
XX diagnosing a disorder related to the female reproductive system,
XX particularly breast and/or ovary cancer. They are also useful in the gene
XX therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
XX agonists and antagonists from the present invention are useful in the
XX diagnosis, treatment and prevention of cancer, immune disorders,
XX cardiovascular disorders, wound healing, neurological diseases and
XX infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
XX used in the exemplification of the present invention.

XX Sequence 128 AA;

Query Match 6.8%; Score 566; DB 21; Length 128;
Best Local Similarity 82.0%; Pred. No. 2.3e-32;
Matches 105; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

QY 1446 DKRGLEVEIIRAGLVYKPGSKTLPAHYKVVYLIDNCVCIAKKTKYAKRTLEPLYOOL 1505

DB 1 dkkgqlvevrlarsltqpgskstppaykvyllengaciacakkkrtlrakltldlyqqs 60

QY 1506 LSPFESPQGRVLOITIVMGDVCRMDKSPMGVAOITLDELISNMVIGMFKLEPPSSLDV 1565

DB 61 lvidespqgkvlygvwdygrmdhkcimgvaqllleeldlssmviygyklltppsslyvdp 120

QY 1566 TSAPLTPRR 1573

DB 121 tltpltrr 128

RESULT 6
AAB40725 standard; Protein: 126 AA.

AC AAB40725;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF489 polypeptide sequence SEQ ID NO:978.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnerary; antiparasitic; antiparkinsonian; neutrotropic; neuroprotective;
XX anticonvulsant; osteopathic; antiallergic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antihemmatic; antithyroid;
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;

KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

OS Homo sapiens.

PN W0200058473-A2.

PD 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US08621.

PR 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 03-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI: 2000-602362/57.

DR N-PSDB; AAC74934.

XX Claim 11; Page 956-957; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX

CC sequences have activities such as: cytostatic; hepatotropic; vulnary;

CC antiproliferative; antiparkinsonian; nootropic; neuroprotective;

CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;

CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;

CC antidiabetic; hypotensive; dermatological; immunosuppressive;

CC antiinflammatory; antibacterial; antiviral; antifungal; antihemetic;

CC antihypertensive; antidiabetic. The sequences can be used for determining

CC the presence of or predisposition to, or preventing or treating

CC pathological conditions associated with an ORFX-associated disorder. The

CC nucleic acids can be used to express ORFX proteins in gene therapy

CC vectors. The proteins and nucleic acids may be used to treat cancers,

CC proliferative disorders, neurodegenerative disorders, osteoarthritis,

CC graft vs host disease, cardiovascular disease, diabetes mellitus,

CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus

CC erythematous, severe combined immunodeficiency (SCID), AIDS, viral,

CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,

CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,

CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance

CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 126 AA;

SO Query Match 5.1%; Score 420; DB 21; Length 126;

Best Local Similarity 65.3%; Pred. No. 4.1e-22;

Matches 81; Conservative 19; Mismatches 24; Indels 0; Gaps 0;

RESULT 7
AAB55848
ID AAB55848 standard; Peptide: 86 AA.
XX
XX AAB55848;
AC
XX
XX 07-MAR-2001 (first entry)
DT
XX
XX PDZ encoded domain #18.

DE

XX Hematopoietic cell; PDZ; PL; autoimmune disease; inflammation;

KW allergy; asthma; multiple sclerosis; cancer; infection.

XX
XX OS Synthetic.

PN W0200069896-A2.

PD 23-NOV-2000.

PF 12-MAY-2000; 2000WO-US13161.

PR 14-MAY-1999; 99US-0134114.

PR 14-MAY-1999; 99US-0134117.

PR 14-MAY-1999; 99US-0134118.

PR 21-OCT-1999; 99US-0160860.

PR 29-OCT-1999; 99US-0162498.

PR 13-DEC-1999; 99US-0170453.

PR 14-JAN-2000; 2000US-0176195.

PR 14-FEB-2000; 2000US-0182296.

PR 11-APR-2000; 2000US-0196460.

PR 11-APR-2000; 2000US-0196527.

XX (ARBO-) ARBOR VITA CORP.

XX Lu PS;

XX WPI: 2001-080245/09.

XX Modulating a biological function of an endothelial cell or

PT hematopoietic cell, useful for treating autoimmune diseases and

PT infectious diseases, by administering an antagonist that inhibits

PT binding between a PDZ protein and a PL protein -

XX disclosure; Page 28-43; 141pp; English.

PS The present invention relates to a new method for modulating a

XX biological function of an endothelial cell or hematopoietic cell. The

CC method involves introducing into a cell, an antagonist that inhibits

CC binding between a PDZ protein and a PL protein. The inhibitor is used

CC to treat a disease mediated by hematopoietic cells, e.g. autoimmune

CC disease. It may also be used to prevent transplantation rejection of

CC a solid organ transplant. The method may also be used in the treatment

CC of inflammation, allergy, inflammatory bowel diseases, ulcerative

CC colitis, ileitis, psoriasis, asthma, atopic dermatitis, autoimmune

CC diseases (e.g. Rheumatoid arthritis, multiple sclerosis,

CC insulin-dependent diabetes, Hashimoto thyroiditis, osteoarthritis,

CC graft rejection, transplantation rejection), atherosclerosis, cancers,

CC infectious diseases, ischemia, vasculitis and Crohn's disease.

XX Sequence 86 AA;

SO Query Match 5.0%; Score 417; DB 22; Length 86;

Best Local Similarity 100.0%; Pred. No. 3.7e-22;

Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 682 RDSGAMGLKVGKMPESGRICAFITVKKGSGLADYVGHLPDGEVLEWNGRLLOGATF 741

DB 1 rdsamglkvvggkmtesgrlcafilkvkgsjadtvgvhlrpgdevlewngrlllgatf 60

QY 742 EEVYNIIESKPEQVELVSR 763

DB 61 eevyniieskpeqvvelvar 82

RESULT	8
ID	AAB57639
AB	AAB57639 standard; Protein; 86 AA.
XX	
AC	AAB57639;
XX	
DT	12-MAR-2001 (first entry)
XX	
DE	PDZ domain.
XX	
KW	Endothelial cell; haematopoietic cell; PDZ domain protein;
KW	PL domain protein; leukocyte activation; synapse formation;
KW	transmembrane neurotransmitter receptor; autoimmune disease;
KW	transplantation rejection; inflammation; allergy;
KW	inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
KW	asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
KW	ischaemia; vasculitis; Crohn's disease.
XX	
OS	Homo sapiens.
XX	
PD	WO20069897-A2.
XX	
PN	23-NOV-2000.
XX	
PF	12-MAY-2000; 2000WO-US13166.
XX	
PR	14-MAY-1999; 99US-0134114.
PR	14-MAY-1999; 99US-0134117.
PR	14-MAY-1999; 99US-0134118.
PR	21-OCT-1999; 99US-0160860.
PR	29-OCT-1999; 99US-0162498.
PR	13-DEC-1999; 99US-0170453.
PR	14-JAN-2000; 2000US-0176195.
PR	14-FEB-2000; 2000US-0182296.
PR	11-APR-2000; 2000US-0196460.
PR	11-APR-2000; 2000US-0196527.
XX	
PA	(ARBO-) ARBOR VITA CORP.
XX	
PI	Lu PS;
XX	
DR	WPI: 2001-025003/03.
XX	
PT	New inhibitors of binding of a PDZ protein and PL protein for
PT	inhibiting T cell-mediated response by hematopoietic cells, or for
PT	treating diseases characterized by inflammatory and humoral immune
PT	responses, e.g. inflammation, cancer
XX	
PS	Disclosure; page 36; 139pp; English.
XX	
CC	The present invention relates to a method for modulating a biological
CC	function of an endothelial cell or haematopoietic cell, comprises
CC	introducing into a cell an antagonist that inhibits binding between a
CC	PDZ domain protein and a PL domain protein to result in inhibition of
CC	leukocyte activation. The present sequence is a PDZ domain. PDZ domains
CC	of proteins are named after three prototypic proteins: PSD95,
CC	Drosophila large disc protein and Zonula Occludin 1 protein. PDZ domain
CC	proteins are involved in synapse formation by organising transmembrane
CC	neurotransmitter receptors through intracellular interactions. The
CC	inhibitors identified by the present invention can be used to treat a
CC	disease mediated by hematopoietic cells, e.g. autoimmune disease,
CC	inflammation, allergy (e.g. drug allergies), inflammatory bowel diseases,
CC	ulcerative colitis, ileitis, psoriasis, respiratory allergic diseases
CC	(e.g. asthma), atopic dermatitis, autoimmune diseases (e.g. rheumatoid
CC	arthritis, multiple sclerosis, insulin-dependent diabetes, Hashimoto
CC	thyroiditis, osteoarthritis), atherosclerosis, cancers, infectious
CC	diseases (e.g. viral infection), ischaemia, vasculitis and Crohn's
CC	disease. The inhibitors can also be used to prevent transplantation
CC	rejection of a solid organ transplant.
XX	
SD	Sequence 86 AA;

Query Match 5.0%; Score 417; DB 22 Length 86;
Best Local Similarity 100.0%; Pred.No.3.7e-22;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps

OY 682 RDSGAMLGKLVVGGKMTESGRCAFIIRKKGSILADTVGHILRPEDEVLFWNGRILOGATP 741
| | | | |
Db 1 rdsamglkvvggkmtesgrlcafiirkvkgsladtvgnlrpgdevlewmgrcllgatf 60
742 EEVYNIIIESKEPQVELVSR 763
| | | | |
Db 61 eevyniiieskepqvelvsir 82

RESULT 9
AAB58049
ID AAB58049 standard; Protein: 86 AA.
XX
AC AAB58049;
XX
DT 12-MAR-2001 (first entry)
XX
DE PDZ domain.
XX
KW Endothelial cell; haematopoietic cell; PDZ domain protein;
KW PL domain protein; leukocyte activation; synapse formation;
KW transmembrane neurotransmitter receptor; autoimmune disease;
KW transplantation rejection; inflammation; allergy;
KW inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
KW asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
KW Ischaemia; vasculitis; Crohn's disease.
OS Homo sapiens.
XX
XX WO200069898-A2.
PN
PD 23-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-USI3205.
XX
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162499.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI: 2001-061214/07.
XX
PT Modulating a biological function of a hematoipoietic cell for treating
PT an allergic response, or diseases mediated by immune system cells,
PT comprises introducing into the cell a PDZ-PL interaction enhancer or
PT inhibitor -
XX
PS Disclosure: Page 38; 143pp; English.
XX
CC The present invention relates to a method for modulating a biological
CC function of an endothelial cell or haematopoietic cell, comprises
CC introducing into a cell an antagonist that inhibits binding between a
CC PDZ domain protein and a PL domain protein to result in inhibition of
CC leukocyte activation. The present sequence is a pdz domain. PDZ domains
CC of proteins are named after three prototypical proteins: PSD95,
CC Drosophila large disc protein and Zonula Occludin 1 protein. PDZ domain
CC proteins are involved in synapse formation by organising transmembrane

CC neurotransmitter receptors through intracellular interactions. The
CC inhibitors identified by the present invention can be used to treat a
CC disease mediated by haematopoietic cells, e.g. autoimmune disease,
CC inflammation, allergy (e.g. drug allergies), inflammatory bowel diseases,
CC ulcerative colitis, ileitis, psoriasis, respiratory allergic diseases
CC (e.g. asthma), atopic dermatitis, autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, insulin-dependent diabetes, Hashimoto
CC thyroiditis, osteoarthritis), atherosclerosis, cancers, infectious
CC diseases (e.g. viral infection), ischaemia, vasculitis and Crohn's
CC disease. The inhibitors can also be used to prevent transplantation
CC rejection of a solid organ transplant.

XX Sequence 86 AA:

Query Match 5.0%; Score 417; DB 22; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.7e-22;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 682 RDSGAMGLKVVGGKMTESGRCAFTTKVKKGLADTVGHLRPGDEVLEWNGRLLOGATF 741
Db 1 rdsamglkvvggkmtesgrlcafttkvkkgladtvghltpgdevlewngrlllgatf 60

QY 742 EEVYNILLESKPQVELVSR 763
Db 61 eevynillesskpqvelvsvr 82

RESULT 10
AAY30795
ID AAY30795 standard; Protein: 1898 AA.

XX AAY30795:

XX 25-NOV-1999 (first entry)

XX A human trichohyalin (TRHY) protein.

XX Human: trichohyalin; TRHY; protein; tissue structure; wound healing;
XX terminally differentiating epidermal tissue; proteinaceous gel;
XX breast implant.

XX Homo sapiens.

XX US5958752-A.

XX 28-SEP-1999.

XX 14-FEB-1997; 97US-0800644.

XX 30-APR-1993; 93US-0056200.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Kim I, Chung S, Park S, Steinert PM, Lee S;

XX WPI: 1999-561041/47.

XX N-PSDB: AA22301.

XX Human trichohyalin useful for forming a proteinaceous gel that promotes
XX wound healing -

XX Disclosure: Fig 3A-W; 126pp; English.

XX The present sequence represents a human trichohyalin (TRHY) protein.
XX The protein is found in terminally differentiating epidermal tissue,
XX and is involved in forming the structural architecture of such
XX tissue. The trichohyalin protein is useful for forming a
XX proteinaceous gel which may then be used for healing wounds, or in
XX breast implants.

XX Sequence 1898 AA;

Query Match 3.7%; Score 307; DB 20; Length 1898;
Best Local Similarity 17.4%; Pred. No. 2.1e-12;
Matches 245; Conservative 213; Mismatches 514; Indels 434; Gaps 53;

QY 19 QPPRPMPDLISHLTEE--RKILAVMDROKKEEKESQVLKIEEKAQPTQMPFSG 76
Db 264 epqqrqlge-----eeeqllkierqlrrrteeqqqlrrrkeq----- 312

QY 77 ITLELVNVLPOOKQPNKEEPQTLHQOFEMKEDQVKMGESQOOQ--EOKGAPTCG 133
Db 313 -----erreqeereqreqqeereqqrlrreqeereqqrlrreqe----- 358

QY 134 ICHKTRFADGCHNCSCYQTFECARCGRSVLSRSKVMWVNCCKROEILLKSAMFWN 193
Db 359 -----trreqll----- 364

QY 194 SGSNTLOOPDQKVPGRGLNEBAPQEKAKLHEDQFOGAPDLSVPAVEKGRHGLTRD 253
Db 365 -----rreqeereqqrlrreqeereqqrlrreqllrreq-----qlrrreqllrreq 412

QY 254 TIKNGGVKHOIASDMSDRKRSVSVDONRYEODEER--EDYSQVPSDGTMPRPSD 312
Db 413 qlrrreqqlrreqqlrreqqlrreqqlrreqqlrreqeeerhkeqreqqrlrreqee 472

QY 313 YAD-----RRSQREPQVFEPEGH--LNYRDSNRGRHRSKEIVYDDEVE 355
Db 473 rrdwlkreeterheqertrkqllkrdgeerterwklleerterqertrreqllrreqee 532

QY 356 SRDE-YERORREVEYOARSRDPMRLARYPVKPOVEQOMRIHAESRARRHERHSDVLA 414
Db 533 rreqrlkrqeereqlrqlrreqqlrr-----eqeereqlkrqeekrlqerreqrl--- 586

QY 415 NAELEDSRISLLRMDRSORSVSERRAMENORSYSMEKTRQAOGSSVPORTSNHSP 474
Db 587 kreqeerrdqlkrqeereqqrlkrqeereqlrr-1kreeverlqeerderlkrreepe 645

QY 475 TPRRSPIPLDRDMRRLADSLRKHNLDRPSSAVRKRKRKMETMLNDSLSQSSSVRP 534
Db 646 eerh-----elkseegeerh-----eqllrreqeereqlkrre-----eeerlegr 690

QY 535 PPRPKSKKGGKMRQVSLSSSEELA-----SPREYSCDVELESSEVSEKGSOKR 588
Db 691 lkreheeer-----regelaeeqegaretrkrlpkw-----qwllesae-----dargsk 737

QY 589 -KRTSEQGVLSDSNTRSERQKKRMVYGGHSLSEDLSESPQIKDSGVDTCSSTLNEHS 647
Db 738 vllapqagaaqbeqqeekrrr-----esellqwe-----eera 772

QY 648 H-----SDKHPTWQPSKDDRLRLILNKRRLKDSVPRDSGAMGLKVYGGKMTESG 701
Db 773 hrqqgeeqtrdftwqwaqeeqsergrqlsar-----prlreqrqrqlraeerqqreqr 827

QY 702 RLCAFTTKVKKGLADTVGHLRPGDEVLEWNGRLLOGATFEVYNILLESKPQVELV 761
Db 828 fl-----peeekeqqgr-----qrrekeqlqf 851

QY 762 SRPIGDIPIRIPDSHQALESSESSSFESQKMDRPSISVSPMSPGMLRDVQFLSQLSIK 821
Db 852 -----lee-----eeqlrrerreqqlqeeedqqlqeeqdeqrrreqqrldq 889

QY 822 LMPDKVGHQLVITLIGANDLPSREDGRPRNPVYKLYFLPDRSDKRRKRYTKTLERPW 881
Db 890 kw-----twql-----eeerkrtrht-lyakpalqeqllrkeqllqeeeeelq 931

QY 882 NQRTIYSPVHREFRERMLETLMDQARVESESEFLLELTALTDDBPHW----- 937
Db 932 ree-----tekrtrreqqeryrr-----eeqllqeeeqllre-----erekrtrreqqeryrrkx 981

QY 938 KLGTHDVSSLPLPRSPYLPRLRLHGESP--TRRLQSRKRLSDSEVSDYDCEDGVGVSD 995
Db 982 klqkqe-----eqllqeepekrtrreqerekkyyree----- 1011


```

Db      627 lre-geaerkp-----vfrkeilkreekkikldrlntvaspckdcgelasiasvgssrps---- 678
QY      809 DVPPQLSGQLSIKIMFKVNGQLVITILGADLPREDGRPNPYK-lyFL-----PDMS 863
Db      679 ---sdlgarly-elagesvenq---evgskkripak-----pqkkqivlddgqper 724
QY      864 DKNRRRTKYVTKLEPPKNNQTFIYSP-----VHRREFRRLLETIMDQAVRE 913
Db      725 dvrykysalrdetpekrsgqekshsvnteekigldhtqytrkqmgerrrykqmemela 784
QY      914 ESEFLG---ELLILEETALL-----DDEPHWYKLTQHDVSLPLPRSPYLP 957
Db      785 ksekfgspkkdvdeyerrslvhevqkppqdvtrddspskkkrmhivdfclckrennys 844
QY      958 RROUH-----GESPTRLQGRSKRISDEVSVDYCEDGVG-----VSDYRHNGDLQS 1005
Db      845 srqisedsertgspsvr-----hgsfthededpigsprllsvkqspkvdekvlypy 894
QY      1006 STLAVPEQVMSNNHCSPSGPHRVDTVIGRTSRWSPPADPPQNVGQHR--GTRATGHYN 1063
Db      895 snltvreeslfnpydsrrregmadmakiklsvlnsedelnwdsqmkqdaqrfdvsvfpn 954
QY      1064 TISRMD--RRRVMD-----HYSRDRDCEADARQPYHRSRSTEQRLERTTSSRSE 1117
Db      955 slkrdsirkrsvrdlepgevpsdsdedge-----hksbprsalyle-----srls- 1002
QY      1118 RPDNNLRMSNPLMTGRAPSPALSRHPRTGSVQTSPTSGTGGRGRLQPLDPKGT 1177
Db      1003 ----fltrdrecklrederlssslern--kysfalaktltpdt-----kal 1044
QY      1178 LER-----SAMDIERNRQMKLNRYKOYAGSDPR-LEDYHSKYSRGMQPHNG 1224
Db      1045 lerksjsssteewnswflawdsrfanfnmkdkevdsapripipwymkklrtldsegg 1104
QY      1225 ADPTVSKSSDSVDVSAVSRSTSSAFSSSYMSVQSEPRPGNKKISYFTSKMONRQMG 1284
Db      1105 md-----dkkedhkeeeqegelfasfrlshsifegdekr-----lqlherkeesdt- 1152
QY      1285 VSGKNLTKSTISGDMCS-----LEKNDGSGS 1311
Db      1153 lsgriylgkqtseganstdsldqepvllfhsrfmeltlrmgqkkkdkpkrkevkgeden 1212
QY      1312 DTAVGALGTSGK----KRSSIGAKMAVIVGLSRKSRASQISQTEGGGKKLRSTYQRET 1367
Db      1213 hpkpesapenkdselkctppsvgppsvtrvcl---esapsalektvg-----dktv 1260
QY      1368 ETGLAVENRMNMTROASRESTDGSNNYS-----SEGNIIFPGVRLAS 1410
Db      1261 eaplvteektvrepativseeakpasepapayeqleqvdlpppadpdkkaammpayveegs 1320
QY      1411 DSQSFDELIDGIPA-----OLVGRQTLATPAM-----GDIQVGMMDKKGL 1451
Db      1321 sgddppjyldakpripagisfgaesnvdpbdpstqplskpdaqkseaneepkaekpdadata 1380
QY      1452 EVEL---IRAGGLVWPKSKTLPAVYVYLLDNGVCIAKKTVAARKTLEFLYQQLLSF 1508
Db      1381 eptdnqkaeapesqpssedle-----vdppvaakdkkpnkskrskrcpvgaaavsi 1432
QY      1509 EESPOGRVLIQIIVWGDYGRMHK-----SFMGVAQIILL-ELLEENMYIGVFKLFPPSS 1561
Db      1433 vekpvtktse-----ridreklkrnsprgeaqkillelmeekit-----rtaskns 1480
QY      1562 LVP-----PTSAPL--TRASQSSLESSTGPGTSRS 1590
Db      1481 aadlhpdpeslipisrtirrvnsvyatmgdheurs 1515

```

```

XX      16-JAN-2001 (first entry)
DT
XX
XX      Human APC protein (splice variant 2).
DE
XX      APC gene; Adenomatous Polyposis Coli gene; human; chromosome 5q21;
XX      familial adenomatous polyposis; FAP locus; Gardner's syndrome; GS;
XX      sporadic tumour; adenoma; carcinoma; cancer; lung; breast; colon; rectum;
XX      bladder; liver; sarcoma; stomach; prostate; leukaemia; lymphoma;
XX      tumour suppressor; anti-APC antibody; detection; diagnosis; prognosis;
XX      genetic predisposition; drug screening; DP2.5; splice variant.
OS      Homo sapiens.
XX
XX      US6114124-A.
XX
XX      05-SEP-2000.
XX
XX      25-MAY-1995; 95US-0450582.
XX
XX      16-JAN-1991; 91GB-0000962.
XX      16-JAN-1991; 91GB-0000963.
XX      16-JAN-1991; 91GB-0000974.
XX      16-JAN-1991; 91GB-0000975.
XX      08-AUG-1991; 91US-0741940.
XX      12-AUG-1994; 94US-0289548.
XX
XX      (ICIL ) IMPERIAL CHEM IND PLC.
XX      (UYJO ) UNIT JOHNS HOPKINS.
XX      (UTAH ) UNIT UTAH.
XX      (CANC-) CANCER INST.
XX
XX      Carlson M, Groden J, Joslyn G, Kinzler K, Markham AF, Anand R;
XX      Albertsen H, White RL, Thilveris A, Nakamura Y, Vogelstein B;
XX      Hedge PJ;
XX      WPI: 2000-565003/52.
XX      N-PSDB; AAA93450.
XX
XX      Detecting Adenomatous Polyposis Coli (APC) protein in a sample for
XX      diagnosing cancers, involves contacting the sample with antibodies that
XX      specifically bind to APC protein and detecting the complex formed -
XX
XX      Example 7; Fig 7A1-7W; 125pp; English.
PS
XX
XX      The invention relates to a novel method for detecting Adenomatous
XX      Polyposis Coli (APC) protein in a sample. The method involves
XX      contacting the sample with antibodies which specifically binds to the
XX      2843 amino acid form of the human APC protein or to a mutant APC
XX      protein, and detecting an APC-antibody complex. Mutations in the APC
XX      gene play a role in tumorigenesis, indicating that it is a tumour
XX      suppressor gene. It is located on chromosome 5q21, which corresponds to
XX      the FAP (familial adenomatous polyposis) locus. FAP is an autosomal
XX      dominant inherited disease in which affected individuals develop
XX      hundreds to thousands of adenomatous polyps in the colon and rectum.
XX      Some of which progress to malignancy. The FAP locus is often found to
XX      be deleted in sporadic (i.e., non-familial) adenomas and carcinomas, and
XX      chromosome 5q deletions have also been observed in tumours of the lung,
XX      breast, colon, rectum, bladder, liver, sarcomas, stomach, and prostate,
XX      and in leukemias and lymphomas. Although the FAP locus contains
XX      several other genes such as FER, TBI, TB2, and MCC, it is thought that
XX      mutations in the APC gene play a key role in the development of FAP and
XX      sporadic tumours. The method is useful for detecting APC protein and its
XX      mutant forms in foetal tissue, placental tissue, amniotic fluid, blood,
XX      serum or a tumour sample. The method is useful for diagnosing or
XX      prognosing neoplastic tissue, for detecting a genetic predisposition to
XX      cancer, for detecting germline and somatic alteration of wild-type APC
XX      genes, and for testing therapeutic agents for the ability to suppress
XX      tumours. The present sequence represents a 2742 amino acid splice
XX      variant of the human APC protein. This variant is less abundant than the
XX      2843 amino acid variant (AAB23011).
XX
XX      Sequence 2742 AA:
SO

```

```

Query Match          3.1%: Score 258.5; DB 21; Length 2742;
Best Local Similarity 17.9%: Pred. No. 9.3e-09;
Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69;

QY 30 SHUTEERKIIIAVMDRCKEKEEKSQVLKIKEHKAOPTQWPFSGITELVNNVLOPOQ 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 911 nhmdndgdelcplnylskysdeq-----Insgrqspqnerwarphkile--deikseq 964

QY 90 KOPNKEKROTKLHQEFEMKEDQKKMGESQO---OOEOKGADPTGCTCKRTFADCGH 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 965 rgrsrgs-----ltypvlytesdtkhikfphfigqe-----cvsyrirganset 1011

QY 147 NCSYQGTFCAGCGGVLSRSKNKVMWVNCILCKQOEILTRKSGAMFYNSGNTLQOPQKV 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1012 n-----rvsnhngingvsgelc-----qeddyeddkpntyseryseeqnee-- 1054

QY 207 PGLRNEEAPQEKKALHBOPOFOGAPGDLSPPAVEKGRAHGLRDTITKNGSVKHQIA 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1055 -----eetrlnyslkyneeckrhvdpidysl-----kya 1083

QY 267 SDMEBDRKRSBVSQRDNRRYEOSEEREDYSQIVPSDGTMPRSPDYADRORSQEPQFYE 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1084 fdipsqkqsfisfksqsgqsakte-----hmsassentstpsnakrqqlhpsaq 1136

QY 327 EFGHLNTRDSNRGRHRSKE-----YIVD-----EDVE 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1137 strsqpqkaatkvsinqetlqtycvedtpricfscsllsisaedeigcqtqead 1196

QY 356 SRDEXEROREEYQARYSDPNLARYPVKPOPEEOMRIHAEVSRARHRRHSDVSLAN 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1197 santlqiaelkekiktriseedpvseavysqbrtkrsrlqgs-slsasesarhkavefs 1255

QY 416 AELERSRISLMDRP-----SKQSVERRAAMENQORSMTREAO-----459
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1256 gakspsksgaqrtpksprehlyvqetplmficrctsvs---sldfestslassvsgsepqsg 1311

QY 460 -----GOSSYQORTSNHSPPTPRRSPIDLPDRMDRADSLRKQHHLDPASAVAKTRERK 513
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1312 mvsqgilspsedlpdsqgqtmprstsktppp-----ppqta-qlkre 1350

QY 514 METMLRNDLSSDSQESVNRPPRPKSKKGGKMRQVSLSSSEEL-----ASTPEXTSC 568
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1351 ---vphkaptakres-----gpkqaavnaavgrvqvipdadtlhhfatestpdcfsc 1401

QY 569 D-----DVL-----BSESVSEKGDQO-KGKRRT--SEOG 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1402 ssslsalsldepfiqkdvellimprvqendngnetesegpkeseenqeakaektlidsek 1461

QY 596 VLSDSN-----TRSEROKRMYYGGHSLSEEDLEMSERPOIKDQSGVDTCSS 639
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1462 lldsdadddeilleecisampkksrkkakpaqtasklpprparkpsqf-----pvyk 1515

QY 640 TTLNEEHSRDKRHYWQPSKDGDRL--IGRIILN-----KRLKDGSVPRDSGAMIGL 690
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1516 llpsqnrlqpqh-vsfirgdpdmprvycevgripinfstslsdltespnelaageyv 1574

QY 691 KVVGGKMTESGRLCAFTTVKKGSLADIVGHLRPGDVLFWNGRLLQGAFTFEVYNNITLE 750
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1575 r--gg--aqsg-----efekrtdlpteg--fstdea-----qgkgtssvltlpeld 1613

QY 751 SKPEQVELVY-----SRIGD-----IPRIPDSTHAOLESSSSSPESOKMRPSTSVT 799
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1614 dhkaeegdliaecinsampkxgshkpfvkkllmdyvq-qasassapnknqlqgkkkkpt 1672

QY 800 SFMSPGMLRDVPOFLSGQSLIKLMPDKVGHQILIVTILGAKDLSPREDGRPNPVIKTYFL 859
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1673 spvkp-----lpqnteirtriv-----rknaacknnlnaerfv 1704

QY 860 PDRSCKNKRRTTVKKTLEPK-----WNQTFYSPVH-----RR 893
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1705 sdnkdkkqnlknkskdfndklpmedrvrgsfafdsphylrplegrypcfsrdsissl 1764

```

```

QY 894 EFRERMEITLMDQARVRESESEFLGEILIELETAALLDDEPHMYKLOTHVDSLPLR-- 951
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1765 dfddddvalsr-ekaeltkakenkesekavtsheteltsngqanqtga--lqkqprnrg 1821

QY 952 PPSVLPKRRQL-----HGESPTRLQR-----SKRISDSFVSDYDCE---987
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1822 pklilqkqstfrpsqskdiptdgaatdekqnfatentpvcfahsslsldsdqgnmk 1881

QY 988 -----DGVGVSVDYRNG-----RLOGSSTLVSPPQVMSNNIC 1020
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1882 enepiketepdpsqgpekskpqasgyapksfhvedtprvcfrmslsslsidseddlqec 1941

QY 1021 SPSGSPHFRVD---VIGTRRSMSPSAPPPQARNVEQGRGTRAGHNTTISRMD-RHRVMD 1076
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1942 issampkkkksprilkygdnekhspp-----rmn-ggllgedtldldkldqrpsehglsdp 1994

QY 1077 HYSDDRDRDCEAAD-----RQ-----PYHRSRSTE 1101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1995 sentdwaikqganaisvslhqaanaaclsrqassdsdlsllskgslsgspfhltpdqe 2054

QY 1102 QRPLL-----ERTTRRSRSERPDTN-----1122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2055 ekpftsnkqprilkygkfstletkkieesekyikgkkykalligkvrnselsgmqkq 2114

QY 1123 -LMRSMPSLMTGRSAPPSPPALSRSHPRTGSV-----QTSSTPTGTRGRGROLPLPP 1174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2115 plqanmpsisrgctmihlpyrvnssstspayekskpplktpskspseg-----qatatsp 2170

QY 1175 KGLERSAMDIEERNRQMLNKYKQVAGSDPRLLEDYHSHKYRSG--WDPHRGAD-----1226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2171 rgakpsvkselsparqet-----sqigqss-----kapsrsgsdstpsrpaqplar 2218

QY 1227 TVSTKSSDSQSDVSAVARTSASRFTSSSTVMSVOSERRGRKRTSVTFKSKMONRQMGVS 1286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2219 plqpsgrnisipgrnqislpnqlsqdlprtspslastksssgqkms-yls--pyrgm--s 2273

QY 1287 GKNLTKSTSIIGDMCSLEKMDCSQSDTAVGALGTSGKRRRSIGAKMVAIVGLSRKRSRA 1346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2274 qgnltkgtglsknaaslrpseasakglnqmnngnank-----velstrmsstk 2322

QY 1347 SOLSQTEGGGKKL-----RSTVORSTETGLAVEMRMNMTROASREST-----1388
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2323 ssgsesdtrseprvlvrgstflkeapspltrkleeasafeslpsrpsprsqgqtpv 2382

QY 1389 -----DGSMSYSYS 1397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2383 lpslpldmsslsths 2397

RESULT 13
AAR63508
ID AAR63508 standard; Protein; 2842 AA.
XX
AC AAR63508;
XX
DT 23-MAY-1995 (first entry)
XX
DE Adenomatous polyposis coli tumour repressor.
XX
KW Adenomatous polyposis coli; tumour repressor; Gardner's syndrome;
KW familial adenomatous polyposis; cancer diagnosis and prognosis;
KW tumorigenesis suppression.
XX
OS Homo sapiens.
XX
PN US5352775-A.
XX
PD 04-OCT-1994.
XX
PF 08-AUG-1991; 91US-0741940.
XX
PR 16-JAN-1991; 91GB-0000962.

```

PR 16-JAN-1991: 91GB-0000963.
PR 16-JAN-1991: 91GB-0000974.
PR 16-JAN-1991: 91GB-0000975.
XX
PA (CANC-) CANCER INST.
PA (ICIL) IMPERIAL CHEM IND.
PA (UYJO) UNIV. JOHNS HOPKINS.
PA (UTAH) UNIV UTAH.
PI
PI Albertsen H, Anand R, Carlson M, Groden J, Hedge PJ,
PI Joslyn G, Kinzler K, Markham AF, Nakamura Y, Thliveris A,
PI Vogelstein B, White RL,
XX
DR WPI: 1994-316233/39.
XX
DR N-PSDB: AA072297.
XX
PT New human adenomatous polyposis coli DNA encoding tumour
PT repressor - and derived primers and probes for diagnosis,
PT prognosis and treatment of cancer
XX
PS Claim 3; Columns 71-84; 113pp; English.
XX
CC AA072297 is a cDNA isolated from the human adenomatous polyposis
CC coli (APC) gene, it encodes the tumour repressors described in AAR63507
CC and AAR63508. Determination of alterations in APC or its expression
CC products, can be used for the diagnosis and prognosis of cancer e.g.
CC colorectal, lung and breast tumours; and for determining predisposition
CC to certain cancers such as familial adenomatous polyposis (FAP) and
CC Gardner's syndrome. The wild type APC gene (or a part of it) can be
CC used therapeutically to restore gene function, while primers and probes
CC derived from the cDNA (AA072333-400 and AA072541-568) can be used to
CC detect mutations. Also APC proteins or analogues can be administered to
CC compensate for a defective gene, and epithelial cells, or transgenic
CC animals carrying a mutated APC allele are useful for detecting
CC therapeutic agents able to suppress tumorigenesis.
XX
SQ Sequence 2842 AA:

Query Match 3.1%, Score 258.5; DB 15; Length 2842;
Best Local Similarity 17.9%, Pred. No. 9.9e-09;
Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69;

QY 30 SHLNEEKRIILAVWDKKEEKEQSVLKIKKEKAQPTQWPFPSGTEIVNNVLOPOQ 89
DB 1011 nhmdndgelctplnyslkysdeq---insgrspsgneraarpkhile--delkysq 1064
QY 90 KQPKNEKEPQTKLHQDFEMWKKEQVKMKGEESQ---QOEQKGDAPTCGICHTKTFADGCGH 146
DB 1065 rgrstrngs-----ttypvtestdckhlkfqhfigqe-----cvspyrstrgaangsel 1111
QY 147 NCSTQCTFCARCGGRVSLRSNKVMVNCMLCKKQOELITKSGAMFYNSGNTLQDPQKV 206
DB 1112 n-----rvgnhbjnqvsgslc-----geddyeddkplnyseryseeqhae-- 1154
QY 207 PGRLENEAPQEKRAKLEHQPOFGAPGLSLVPAVEKGRAGHGTRODTITKNSGVAKHQA 266
DB 1155 -----eerpnylsikyeeekrhvqdpidysl-----kya 1183
QY 267 SDMPSDRRKSPSVSRDQNRRYEQQSEEREDYSQVPSDGTMPSPDYADRRSQREFQFE 326
DB 1184 tdlpsqsgsfstfksqsgskste-----hmssseentstpsnakrrqnlhpsaq 1236
QY 327 EPGHLNRYDSNRGRHSKE---YIVD-----EDVE 355
DB 1237 srsgqpqkaactkvsslnqetqlycvedtpicfscsissaeedlgcngtqead 1296
QY 356 SRDEEQRGRREVEYQARYRSDPNLARYPKQPYEQEOMRIHAEVSAARERHRSDSLAN 415
DB 1297 santqtaieikqigtrtaedpvyevpavsqhprtkstrlqgs--slssesarhkavefs 1355
QY 416 AELDESRISLRLMDP-----SRORSVSERRAAMENORSYSMEPRTEAO--- 459
DB 1356 gakspsksqgqtpksphehyqetplmfsrcsvs-----slsfesrslasevqgpcsg 1411
QY 460 -----GQSSVPRQRTSNHSPPTPRPRSPRIPLDRPDMRRADSLRKQNHLDPSAVKRTKREK 513
DB 1412 mvsglispstldpsrpgqtmprstsktppp-----ppqta--qtkre- 1450
QY 514 METMLRNDLSLSDQSESVRPPPHKSKKCGKMRQVSLSSSEEL-----ASTPEYNSC 568
DB 1451 ---vpknkaptaeeres-----gpkgaavnaavqrvqylpdadlllfnafestpddgfscc 1501
QY 569 D-----DVEL-----ESESVEKGDQ--KGKRRT--SEOG 595
DB 1502 sssisalsldepfikqdvrltimpvgndngnetesepkxeneqgkeeketldsekd 1561
QY 596 VLSDSN-----TSEKQKRMYYGCHSLLEDLEMSQIQINDSGVDTGSS 639
DB 1562 lldsdoddieilleecisamptksrkkpqaasklppvarkpsql-----pyvk 1615
QY 640 TTLNEESHSDHPVTPQPSKDGRL--IGRIILN-----KRUKDSVPRDSGAMGL 690
DB 1616 llpsgnrlgpqkh-vstftpgddmpryvcvegtplnfstatslstdtiesppnelaageyv 1674
QY 691 KVGGMKTESGRLCAFTTKVKKGLADTVGHLRPGDEVLEWNGRLLGATEEYVNIILE 750
DB 1675 r--gg--agsg-----efekrttlpteg--rstdae-----ggktsavtllpeid 1713
QY 751 SKRPEQVELVY-----SRPID-----IPRIPDTHAQLESSESSFSQKMDRSISVT 799
DB 1714 dnkaeeqdillaeclinsampkqsnkpfvkkimdvqg--qaasassaaphknqldgkkrpt 1772
QY 800 SPMSPGMLRDVPOLSQSLIKLWFDKQGHQILYITLAKADLPSEDDRPRPYVKIYFL 859
DB 1773 spvkr-----ipqnteyrlv-----tknadsknlnhaevf 1804
QY 860 PDRSDKNNKRRRTVKTKLEPK-----WNQFTIYSPVH-----RR 893
DB 1805 sdnokskkqnlknskndfndklnmedrvrgsfadspbhylptegtrpycfersndsl 1864
QY 894 EFERMLEITLMDQARVRESESEFLGELILELTAALLDDEPHWYLGQNDVSLPLR-- 951
DB 1865 dfddddvdlsr-ekaeltrkakenkeseeakvsheltnsgsaanktga--lakgplnrgq 1921
QY 952 PSPYLPRL-----HGESPTRLQ-----SKRISDEVSVDYDCE----- 987
DB 1922 kpriilqgstfpqsgskdtpdgaatdeklnfalentpvcfshmslslsldidennk 1981
QY 988 -----DGVGVVSDYRHNQ-----RDLOSSTLVSPQVMSNHC 1020
DB 1982 enepiketepdsggepskpgasgyapksflivedprvcfsnsslsisdseddllqec 2041
QY 1021 SPSSGSPHRVD---YIGTRSWSPSAPPPQQRVNEQCHKSTRATGHTNTISRMD--RIRVMD 1076
DB 2042 issampkklkpsrlkgnedkhs-----rnm-gqilgedltdldkqrpdsenglspd 2094
QY 1077 HYSRDDRDRCFAAD-----RO-----PYHRSRST 1101
DB 2095 sentdwkaigegansivslngaaaaacsltrgaasdsdlslsksgislsfpnltpde 2154
QY 1102 QRPLL-----EKTTTTRSRSSERPDTN----- 1122
DB 2155 ekpftsmkgprllkpgketletkkleeskgikgkkyvkslltgkvrnselsgmqk 2214
QY 1123 -LMSRMSLMTGRSAPSPALSRSHPRGSV-----QTSSTSTPGGRKGRQLPOLP 1174
DB 2215 plqanmpsisrgrtmllhpgvrnssstspvskkypplktpkaskspseg---qatatsp 2270
QY 1175 KGLTERSAMDIIEERNQMLKMKYQOVAGSDPRLLEDYDSKVRSG---WDPHRGAD----- 1226
DB 2271 rgakpvtkselsparqt-----sqiggs-----kapsrsgsdstprpaqplsr 2318
QY 1227 TVSTKSSDSQSDVSAVSRSSASRFSSTYSWVSEPRPGNRKRTSVFTSKMQRMGVS 1286
DB 2319 plqspgrnslspgrnslspplklsqllptsspsatasksgsgkms-yts--pqrqm--s 2373

DB 1356 gakspsksqgqtpksphehyqetplmfsrcsvs-----slsfesrslasevqgpcsg 1411
QY 460 -----GQSSVPRQRTSNHSPPTPRPRSPRIPLDRPDMRRADSLRKQNHLDPSAVKRTKREK 513
DB 1412 mvsglispstldpsrpgqtmprstsktppp-----ppqta--qtkre- 1450
QY 514 METMLRNDLSLSDQSESVRPPPHKSKKCGKMRQVSLSSSEEL-----ASTPEYNSC 568
DB 1451 ---vpknkaptaeeres-----gpkgaavnaavqrvqylpdadlllfnafestpddgfscc 1501
QY 569 D-----DVEL-----ESESVEKGDQ--KGKRRT--SEOG 595
DB 1502 sssisalsldepfikqdvrltimpvgndngnetesepkxeneqgkeeketldsekd 1561
QY 596 VLSDSN-----TSEKQKRMYYGCHSLLEDLEMSQIQINDSGVDTGSS 639
DB 1562 lldsdoddieilleecisamptksrkkpqaasklppvarkpsql-----pyvk 1615
QY 640 TTLNEESHSDHPVTPQPSKDGRL--IGRIILN-----KRUKDSVPRDSGAMGL 690
DB 1616 llpsgnrlgpqkh-vstftpgddmpryvcvegtplnfstatslstdtiesppnelaageyv 1674
QY 691 KVGGMKTESGRLCAFTTKVKKGLADTVGHLRPGDEVLEWNGRLLGATEEYVNIILE 750
DB 1675 r--gg--agsg-----efekrttlpteg--rstdae-----ggktsavtllpeid 1713
QY 751 SKRPEQVELVY-----SRPID-----IPRIPDTHAQLESSESSFSQKMDRSISVT 799
DB 1714 dnkaeeqdillaeclinsampkqsnkpfvkkimdvqg--qaasassaaphknqldgkkrpt 1772
QY 800 SPMSPGMLRDVPOLSQSLIKLWFDKQGHQILYITLAKADLPSEDDRPRPYVKIYFL 859
DB 1773 spvkr-----ipqnteyrlv-----tknadsknlnhaevf 1804
QY 860 PDRSDKNNKRRRTVKTKLEPK-----WNQFTIYSPVH-----RR 893
DB 1805 sdnokskkqnlknskndfndklnmedrvrgsfadspbhylptegtrpycfersndsl 1864
QY 894 EFERMLEITLMDQARVRESESEFLGELILELTAALLDDEPHWYLGQNDVSLPLR-- 951
DB 1865 dfddddvdlsr-ekaeltrkakenkeseeakvsheltnsgsaanktga--lakgplnrgq 1921
QY 952 PSPYLPRL-----HGESPTRLQ-----SKRISDEVSVDYDCE----- 987
DB 1922 kpriilqgstfpqsgskdtpdgaatdeklnfalentpvcfshmslslsldidennk 1981
QY 988 -----DGVGVVSDYRHNQ-----RDLOSSTLVSPQVMSNHC 1020
DB 1982 enepiketepdsggepskpgasgyapksflivedprvcfsnsslsisdseddllqec 2041
QY 1021 SPSSGSPHRVD---YIGTRSWSPSAPPPQQRVNEQCHKSTRATGHTNTISRMD--RIRVMD 1076
DB 2042 issampkklkpsrlkgnedkhs-----rnm-gqilgedltdldkqrpdsenglspd 2094
QY 1077 HYSRDDRDRCFAAD-----RO-----PYHRSRST 1101
DB 2095 sentdwkaigegansivslngaaaaacsltrgaasdsdlslsksgislsfpnltpde 2154
QY 1102 QRPLL-----EKTTTTRSRSSERPDTN----- 1122
DB 2155 ekpftsmkgprllkpgketletkkleeskgikgkkyvkslltgkvrnselsgmqk 2214
QY 1123 -LMSRMSLMTGRSAPSPALSRSHPRGSV-----QTSSTSTPGGRKGRQLPOLP 1174
DB 2215 plqanmpsisrgrtmllhpgvrnssstspvskkypplktpkaskspseg---qatatsp 2270
QY 1175 KGLTERSAMDIIEERNQMLKMKYQOVAGSDPRLLEDYDSKVRSG---WDPHRGAD----- 1226
DB 2271 rgakpvtkselsparqt-----sqiggs-----kapsrsgsdstprpaqplsr 2318
QY 1227 TVSTKSSDSQSDVSAVSRSSASRFSSTYSWVSEPRPGNRKRTSVFTSKMQRMGVS 1286
DB 2319 plqspgrnslspgrnslspplklsqllptsspsatasksgsgkms-yts--pqrqm--s 2373

Db 1676 r--g--agsg-----efekrtiptleg--rsidea-----ggktsvtlipeld 1714
QY 751 SKPEQVELV-----SRPIGD-----IPRIPOSTHAOLESSSSFEQKMDRPSISVT 799
Db 1715 dnkaeegdllaecinsampkqkshkpfvkkimdvq--qasassapnknqldgkkkkpt 1773
QY 800 SPMSPGMLRDVPOPLSGLSIRKIMFDKVGHOLIYIILGAKDLPSREDGRPRNPYKIYFL 859
Db 1774 spvkr-----ipqnteyleftv-----fknaadsknlnaervf 1805
QY 860 PDRSDKNKRRRTKTKTLEPK-----WNQTFIYSPVH-----RR 893
Db 1806 sdnkdsdkgnknkskdndklpmedrvgsfafdsphhyplegtpcfsrndsissl 1865
QY 894 EFRERMLEITLMDQARVREESEFLGELLLEFALLDDEPHWYKLOTHDVSSLPLPR-- 951
Db 1866 dfddddvdlstr--ekaelrkaekenkesaekvtsheltensqgsanktqa--lakpinqrl 1922
QY 952 PSPYLPRRQL-----HGESPTRRLQR-----SKRISDSEVSDYDCE---- 987
Db 1923 pkpiliqyqstfpgsskdiqdrgaadexlqnfalientpvcfshmslsslsidqennmk 1982
QY 988 -----DGVCVVDYRHHG-----RDLQSTLSVPEQVMSNHC 1020
Db 1983 enepiketepdpdsgqepskpqasgyapksfhwedtpvcfsrnsjsslsidseddlqec 2042
QY 1021 SPSSGSPHREV--VIGTRMSPPSAPPORVNEQGHGCTRATGHYNTISRMD--RHRVMD 1076
Db 2043 issampkkkksrllkqdnkhspp-----rnm-ggllgedlcldkldqrpdsahglspd 2095
QY 1077 HYSDDRDCFEAD-----RQ-----PYHRSSTE 1101
Db 2096 sentfawkaigeanisvslhqaanaaclsrqassdsdlsllksqslsgspflhlpdqe 2155
QY 1102 QRPL-----ERTTRESRSERPDTN----- 1122
Db 2156 ekpftsnkqprllkpgekstletkkleesekgikgkvykalltgkvrnsneisqmkq 2215
QY 1123 -LMRSMPLMTGRSAPSPALSRSHPRTGSV-----OTSPSTPGTGRGRQPLPOLPP 1174
Db 2216 plqanmpisrgrtmihlpyvrnsstspvskkqprllktpaskspseg---qtattp 2271
QY 1175 KGTLEKAMDIEERKQKMLNKYQVAGSDPRLEODYHSKYRSG--WDPHRCAD----- 1226
Db 2272 rgakpsvkselsparqt-----sqlygss-----kapsrsgsrdsctprpaqplsr 2319
QY 1227 TVSTRSSDSDDVSAVARTSASRSFSSSTSYMVOSEPRPGNKKISVFTSKMONQMGVS 1286
Db 2320 plqspgrnsispgrrngisppnklslqllprtspsstasksgskms-yls--pgrrqm--s 2374
QY 1287 GKNLTKSTISIGDMCSLEKNDGSQSDTAVGALGTSCKRRRSIGAKMVAIVGLSKRSRGA 1346
Db 2375 qgnllkqclqisknaasiprsesaaskylngmnngankk-----velstmsstk 2423
QY 1347 SOLSTEGEGGKKL-----RSTVORSTETGLAVEMRWMTROASREST----- 1388
Db 2424 ssgesdtrseprvlvrgstflkeapspllrklleesastfsspsrpaaplrsgaqtpv 2483
QY 1389 -----DGSMSNYS 1397
Db 2484 lpslpldmslstrhs 2498

Search completed: November 21, 2001, 16:05:29
Job time: 320 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:06:19 ; Search time 16.56 Seconds
(without alignments) 3289.023 Million cell updates/sec

Title: US-09-617-099b-1

Sequence: 1 MSAPLPGPRGPRPAPASQP.....TRRASQSLESNGPSYSRS 1590

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SWISSProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	338	4.1	1021	1 YPT7_CAEEL	P41885 caenorhabdi
2	307	3.7	1898	1 TRHY_HUMAN	Q07283 homo sapien
3	262.5	3.2	684	1 RP3A_RAT	P47709 rattus norv
4	261.5	3.2	704	1 RP3A_BOVIN	Q06846 bos taurus
5	258.5	3.1	2843	1 APC_HUMAN	P25054 homo sapien
6	255.5	3.1	2349	1 TPR_HUMAN	P12270 homo sapien
7	255	3.1	2842	1 APC_RAT	P70478 rattus norv
8	254.5	3.1	2845	1 APC_MOUSE	O61315 mus musculu
9	254	3.1	2468	1 MAPB_HUMAN	P46821 homo sapien
10	253	3.1	2440	1 NCR1_HUMAN	O75376 homo sapien
11	248.3	3.0	1453	1 NKCR_MOUSE	P30415 mus musculu
12	244	2.9	1462	1 NKCR_HUMAN	P30414 homo sapien
13	242	2.9	1861	1 MAP2_RAT	P15146 rattus norv
14	241	2.9	1549	1 TRHY_SHEEP	P22793 ovis aries
15	239.5	2.9	2476	1 ATRX_MOUSE	O61687 mus musculu
16	233.5	2.8	1407	1 TRHY_RABIT	P37709 oryctolagus
17	230.5	2.8	2453	1 NCR1_MOUSE	Q60974 mus musculu
18	224	2.7	606	1 RP3A_MOUSE	P47708 mus musculu
19	223	2.7	2142	1 BAT2_HUMAN	P48634 homo sapien
20	221.5	2.7	1828	1 MAP2_MOUSE	P20357 mus musculu
21	221	2.7	2375	1 ATRX_HUMAN	P46100 homo sapien
22	218	2.6	1781	1 AKAC_HUMAN	O02952 homo sapien
23	217.5	2.6	2774	1 MAPA_RAT	P34926 rattus norv
24	216	2.6	1815	1 UN13_CAEEL	P27173 caenorhabdi
25	214	2.6	1411	1 TCOF_HUMAN	O13428 homo sapien
26	208.5	2.5	1395	1 SP41_YEAST	P38904 saccharomyc
27	208	2.5	2472	1 NCR2_MOUSE	O94442 mus musculu
28	207.5	2.5	1490	1 CRK7_HUMAN	O94444 homo sapien
29	207.5	2.5	2464	1 MAPB_MOUSE	P14873 mus musculu
30	206	2.5	1805	1 NEST_RAT	P21263 rattus norv
31	206	2.5	2805	1 MAPA_HUMAN	P78559 homo sapien
32	205	2.5	1957	1 YD86_SCHPO	Q10411 schizosacch
33	205	2.5	3924	1 ANK2_HUMAN	Q01484 homo sapien

34	203.5	2.5	1676	1 APSA_EMBNI	Q00083 emericeila
35	200	2.4	927	1 NEDA_HUMAN	P46934 homo sapien
36	199.5	2.4	568	1 INVO_RAT	P48998 rattus norv
37	199.5	2.4	3038	1 TRIO_HUMAN	O75962 homo sapien
38	198	2.4	1658	1 YME7_YEAST	O03651 saccharomyc
39	198	2.4	2004	1 MOZ_HUMAN	O92794 homo sapien
40	197.5	2.4	3969	1 HRX_HUMAN	O03164 homo sapien
41	197	2.4	2150	1 SDC3_CAEEL	P34706 caenorhabdi
42	196.5	2.4	1189	1 XJH6_YEAST	P47035 saccharomyc
43	196	2.4	388	1 SYT3_RAT	P40748 rattus norv
44	195.5	2.4	3130	1 DPO2_HUMAN	O60673 homo sapien
45	195	2.4	1075	1 NF33_HUMAN	Q12968 homo sapien

ALIGNMENTS

RESULT	ID	YP77_CAEEL	STANDARD	PRT	1021 AA.
AC	P41885	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	01-FEB-1996 (Rel. 33, Last annotation update)				
DE	HYPOTHETICAL 111.7 KDA PROTEIN F37A4.7 IN CHROMOSOME III.				
GN	F37A4.7				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;				
OC	Rhabditidae; Peloderinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RA	Fulton L., Waterston R.;				
RL	Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.				
CC	-1- SIMILARITY: CONTAINS 2 C2 DOMAINS.				
CC	-1- SIMILARITY: SOME, TO YEAST YNL087W.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: U00032; AAA50635.1; -				
DR	HSSP: P21707; IRSY.				
DR	Wormpep: F37A4.7; CE00705.				
DR	InterPro: IPR000008; -				
DR	InterPro: IPR001565; -				
DR	Pfam: PF00168; C2; 2.				
DR	PRINTS: PR00360; C2DOMAIN.				
DR	PRINTS: PR00399; SYNAPTOGAMN.				
DR	PROSITE: PS00499; C2_DOMAIN_1; 2.				
DR	PROSITE: PS50004; C2_DOMAIN_2; 2.				
KW	Hypothetical Protein; Repeat.				
FT	DOMAIN 758 .. 847 C2 DOMAIN.				
FT	DOMAIN 898 .. 987 C2 DOMAIN.				
SO	SEQUENCE 1021 AA; 111749 MW; 947C38B7C37BA462 CRC64;				
Query Match 4.1%; Score 338; DB 1; Length 1021;					
Best local similarity 20.9%; Pred. No. 5.1e-09;					
Matches 197; Conservative 139; Mismatches 342; Indels 266; Gaps 43;					
OY	90 KQNEKEPQKHLQGEPMYKQYKMM---GEESQOQDEOK-----GDAP 130				
DB	114 RSTNSAOTGSGITPAE--OEHTOKVLAKAEESKSKQDRIGKAVDRLEKRRRATGCGV 171				
OY	131 T-CGICHTKTF---ADGCCHNSYCQTKFCAR-CGGR---VSLRSNKV---MMVCNLCRK 179				
DB	172 THGLLCH-TFEGILLASVYAMCVDCRKYVCGRRNGVETTDVNGTGTGVEIVFLCKICSE 230				

QY	180	QOELL-ITKSGAMTYNSSSNTLQORDQKVRGL---	ENEAPQCKKKKLLHDPQFOGAPG-	23	4
QY	231	AREVLLKSGAMETKE-MPEQQRDDLPITYVVTNGLPLNASSA---- <td>ATPLSGTPOG</td> <td>285</td> <td></td>	ATPLSGTPOG	285	
QY	235	-----DLSEVPAVE-----	KGRAHGLRDTOTLNGSGVKKHQTASDMPSDKKRS-----	276	
Db	286	AGQPMPTMPTSSQOMTTPKMASPGVCSNPGLOMNGP-----	TSPLPNCGRMTWGCHGIE	341	
QY	277	-PVSYSRQNKRYEDESEREDYSQYV----	SDQTMPSBDYADKRSQRPQYEEBGLH	331	
Db	342	FPRSSSGESFVQGYPRRALNKKTPGSGTSAATSPAPPTSTTPSRRE-----	A	392	
QY	332	NYRDSNRGRJHNSKEYLVDEDEVSREYRORREEEYQARYKSDPNILARYPVQPOYEE	391		
Db	393	NMEFSNHTAHANRLKSTDDDDSSSESPTSRIS-----	PHSL-----	434	
QY	392	QMRTHAEVSPRANHERRHSDVSLANAELEDISRILKMDRESRQORSYERRAMENORSYS	451		
Db	435	-----ATPSSVAHDTC-HPSLPADTRSDSGVQSDHNSPOQS-----	GLTCS	478	
QY	452	MEKTRKQOGSSYQQRKTSNHS-PTTPRKSITPLDRPMKRAJDSUKRONKQDSSAARKTK	510		
Db	479	SSSLTPLOOQASH-----DHHSGGCTPRR-----	ISNDP-----RTTS	511	
QY	511	REKMETMLNDSLSDQSESVRPPRRPHKSKKGKKKQVSLSSSEELASTPEYTSQSD	570		
Db	512	RV-----AGASAGTSLVTPRP-----	ISSKTSPPNCSSPLNV-----	545	
QY	571	VELESSEVSEKDGSKQKRRKT-SBQVGLSDSNTRFSERQOKRMUYGGHSLDELEWSEPOI	629		
Db	546	--MEKSSASTASSGGNRKRGSAEPLVNLNHNAMHNNQ-----	NHNDINKKLISQTSR	596	
QY	630	KDSGVDCSSITLLEHSHSKNHNVTQPSKDGRLGRLIKRLKDGSVPRDSGAMLG	689		
Db	597	AESPLAASSSFLSSPPDDTKQKN-----RRDDGGRVNSJQLRTSLDVAAP-----	643		
QY	690	LKVVGKMTESGRIACAEITVKKKKSLADTVGHILPRGDEVEMNORLLQOGATFEVYNIIL	749		
Db	644	-----VARISKNN-----	GH-----IV	655	
QY	750	ESKREPOV--ELVVSRIPIGDIPRIPIPSH--AOLESSSSSFSQKMDRPSISVTSMPSP	804		
Db	656	SSEPTSSSTSNQNTSVPIPVVPEEKEKATITASTESASEGVAIEPLDIDENLEP	715		
QY	805	-----GMLROY-----	POPLSGOLSIKMDKYGHOILVILAKDLPSPEDG	847	
Db	716	KHASRRRGDVGVRGFCLLCFSRKORSLSGTTLLTYHSADKIKLMLIRAKNLIKADSN	775		
QY	848	RPRRPVYVYIPLPRSDKNKRRRTVTKLTLEPKKNQFIYSPVRRFERRERMLEITLMDQ	907		
Db	776	GFSDPYVAFHLLPONTKATKILSTIKTLINPENNEMSYGITEDEKKEKILRVTVLDR	835		
QY	908	ARVRESESELGELLIEETALLDDEPHMYKLGTHDVSLPLRP	951		
Db	836	DRI--GSDPLGETRIALK-KLINENMKKEFTLYLE--SALVPVQ	873		
RESULT 2					
TRHY_HUMAN	TRHY_HUMAN	STANDARD;	PRT;	1898	AA.
AC	007283;				
DT	01-OCT-1994 (Rel. 30, Created)				
DT	01-OCT-1994 (Rel. 30, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	TRICHOHYALIN.				
GN	THH OR TRHY OR THL.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RP	[1]				

```

RA MEDLINE=93280194; PubMed=7685034;
RX Lee S.-C., Kim I.-G., Marekov L.N., O'Keefe E.J., Parry D.A.D.,
RA Steinhert P.M.;
RT 'The structure of human trichohyalin. Potential multiple roles as a
RT functional EF-hand-like calcium-binding protein, a cornified cell
RT envelope precursor, and an intermediate filament-associated (cross-
RT linking) protein.';
RL J. Biol. Chem. 268:12164-12176(1993).
RN [2]
RP SEQUENCE OF 1731-1898 FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=93315897; PubMed=7686953;
RA O'Keefe E.J., Hamilton E.H., Lee S.-C., Steinhert P.M.;
RT "trichohyalin: a structural protein of hair, tongue, nail, and
RT epidermis.";
RL J. Invest. Dermatol. 101:655-715(1993).
CC -1- FUNCTION: INTERMEDIATE FILAMENT-ASSOCIATED PROTEIN THAT ASSOCIATES
CC IN REGULAR ARRAYS WITH KERATIN INTERMEDIATE FILAMENTS (KIF) OF THE
CC INNER ROOT SHEATH CELLS OF THE HAIR FOLLICLE AND THE GRANULAR
CC LAYER OF THE EPIDERMIS. IT LATER BECOMES CROSS-LINKED TO KIF BY
CC ISODIPEPTIDE BONDS. IT MAY SERVE AS SCAFFOLD PROTEIN, TOGETHER
CC WITH INVOLUCRIN, IN THE ORGANIZATION OF THE CELL ENVELOPE OR EVEN
CC ANCHOR THE CELL ENVELOPE TO THE KIF NETWORK. IT MAY BE INVOLVED IN
CC ITS OWN CALCIUM-DEPENDENT POSTSYNTHETIC PROCESSING DURING TERMINAL
CC DIFFERENTIATION.
CC -1- SUBUNIT: MONOMER (PROBABLE).
CC -1- TISSUE SPECIFICITY: FOUND IN THE HARD KERATINIZING TISSUES SUCH AS
CC THE INNER ROOT SHEATH (IRS) OF HAIR FOLLICLES AND MEDULLA, AND IN
CC THE FLILIFORM PAVILLAE OF DORSAL TONGUE EPITHELIUM (PROBABLE).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING LATE DIFFERENTIATION OF
CC THE EPIDERMIS.
CC -1- DOMAIN: CONSISTS OF NINE DOMAINS. DOMAIN 1 CONTAINS TWO EF-HAND
CC CALCIUM-BINDING DOMAINS. DOMAINS 2-4, 6, AND 8 ARE ALMOST
CC ENTIRELY ALPHA-HELICAL, CONFIGURED AS A SERIES OF PEPTIDE REPEATS
CC OF VARYING REGULARITY, AND ARE THOUGHT TO FORM A SINGLE-STRANDED
CC ALPHA-HELICAL ROD STABILIZED BY IONIC INTERACTIONS. DOMAIN 6 IS
CC THE MOST REGULAR AND MAY BIND KIF DIRECTLY BY IONIC INTERACTIONS.
CC DOMAINS 5 AND 7 ARE LESS WELL ORGANIZED AND MAY INDUCE FOLDS IN
CC THE MOLECULE. DOMAIN 9 CONTAINS THE C-TERMINUS, CONSERVED AMONG
CC DIFFERENT SPECIES.
CC -1- PTM: KNOWN SUBSTRATE OF TRANSGLUTAMINASE. SOME 200 ARGININES ARE
CC PROBABLY CONVERTED TO CITRULLINES BY PEPTIDYLARGININE DEIMIDASE.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE S-100
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L09190; AAA65582.1; -.
CC PIR: A45973; A45973.
CC DR HSSP: P02633; 1BOC.
CC DR MIM: 190370; -.
CC DR InterPro: IPR001751; -.
CC DR InterPro: IPR002048; -.
CC DR Pfam: PF01023; S.100; 1.
CC DR Pfam: PF00036; eHand; 1.
CC DR PROSITE: PS00018; EF_HAND; 1.
CC DR PROSITE: PS00303; S100_CABP; 1.
CC DR Repeat; Calcium-binding.
CC FT DOMAIN 1 91 S-100 LIKE.
CC FT CA_BIND 22 33 SITE II (LOW AFFINITY) (POTENTIAL).
CC FT CA_BIND 62 73 SITE I (HIGH AFFINITY) (POTENTIAL).
CC FT DOMAIN 314 390 6 X 13 AA TANDM REPEATS OF
CC FT REPEAT 314 326 R-R-E-Q-E-E-R-R-E-Q-Q-L.
CC FT REPEAT 327 339 1-1 (APPROXIMATE).
CC FT REPEAT 340 351 1-2 (APPROXIMATE).
CC FT REPEAT 352 364 1-3 (APPROXIMATE).
CC FT REPEAT 352 364 1-4.

```

FT	REPEAT	365	377	1-5.
FT	REPEAT	378	390	1-6.
FT	DOMAIN	391	444	9 X 6 AA TANDEM REPEATS OF R-R-E-Q-Q-L.
FT	REPEAT	391	396	2-1.
FT	REPEAT	397	402	2-2.
FT	REPEAT	403	408	2-3.
FT	REPEAT	409	414	2-4.
FT	REPEAT	415	420	2-5.
FT	REPEAT	421	426	2-6.
FT	REPEAT	427	432	2-7.
FT	REPEAT	433	438	2-8.
FT	REPEAT	439	444	2-9.
FT	DOMAIN	444	702	9 X 28 AA APPROXIMATE TANDEM REPEATS.
FT	DOMAIN	923	1162	8 X 30 AA TANDEM REPEATS.
FT	REPEAT	923	952	4-1.
FT	REPEAT	953	982	4-2.
FT	REPEAT	983	1012	4-3.
FT	REPEAT	1013	1042	4-4.
FT	REPEAT	1043	1072	4-5.
FT	REPEAT	1073	1102	4-6.
FT	REPEAT	1103	1132	4-7.
FT	REPEAT	1133	1162	4-8.
FT	DOMAIN	1250	1849	23 X 26 AA APPROXIMATE TANDEM REPEATS.
FT	CONFLICT	1752	1752	F -> L (IN REF. 2).
FT	CONFLICT	1794	1801	OEEDROYR -> RSETGSG (IN REF. 2).
FT	CONFLICT	1857	1857	Q -> K (IN REF. 2).
FT	CONFLICT	1880	1880	V -> G (IN REF. 2).
Q	SEQUENCE	1898 AA:	247219 MW:	A74B5947FB62E31D CRC64;

Query Match	3.7%;	Score 307;	DB 1;	Length 1898;
Best Local Similarity	17.4%;	Pred. No. 3.3e-07;		
Match 245;	Conservative 213;	Mismatches 514;	Indels 434;	Gaps 53;

QY	19	OPRORRPMRLSHLTTEE--RKITLWMDOKKEEKEOSVLTKENKQAPOMPFPSG	76
Db	264	EPORORELDE---EEEOJLKEORELREEEEOOOLRLREOOLRLKOE----	312
QY	77	ITELVNVNLOPOOKOPMEKBPOTLHOOFMYKEOVKMGEOOOO---EOKDAPYCS	133
Db	313	-----EKEOEOERREDOERREDOERREDOOLRLDEORENREOOLRLDEORE	358
QY	134	ICHKTFADCGHNCYSYCQTFKCARCGSVLSNKMVMYCNLCRKOOELITKSGAFYN	193
Db	359	-----RRROOL-----	364
QY	194	SGSUTLOOPOOKVYRGRLNEDROEKKAKLHEDORFOGADOLSVRAVEKGRANGLTOD	253
Db	365	-----RRROEOERREOOLRLREOERREOOLRLREOOLREO-----OLRREOOLREQ	412
QY	254	TIKNGSGVKNHIASDMSDRKSPSVSRDONRKYEOSEER-EDYSQVVPDSGYMPSPD	312
Db	413	OLRREOOLRLREOOLRLREOOLRLREOOLRLREOOLREOEEENHEOKNEBKEERREOOLRLKROEE	472
QY	313	YAD-----RRSOREROFYEERGN---LWYRDSNRGRNHSKEYIVDDEVE	355
Db	473	RDMYKREETERERHEERBRKQOLKRDDEERREBEVTKLEERERREOERREOOLRLRROEE	532
QY	356	SRDE-YERQOREEETQARYSDRYPLARYKPRODYEBOMKINAEVSAKNERKHSVSLA	414
Db	533	RRROJLKRROEEERLQOLRLREOOLRLR---EOERLLOOLKREBEKLEDOERROJL---	586
QY	415	NAELDESRITSLYMDRPSROSVSEBRAAMENOSYSMERTREAOOGSSYPORTSNHSP	474
Db	587	KREOERKROOLKKEERROOLKREOERBEJQR-LKREVEYRLDEORDERLKKREPE	645
QY	475	TPRRSPILDRPDMRADSLRKONHLDPSSAVRKTREKMETYLRLNDSLSDOSESVRRP	534
Db	646	EERH-----ELKSEDOERH---DOLRREOERREOOLRLKRE-----EERHLEJQR	690
QY	535	PRPHKSKGGMKOVYLSSEBEA-----STREYTSODVLESESVSEKDSOKG	588
Db	691	LKRHEEER---REOEIAEEOOAKERIKSLRPK---OMOLEEA-----DAOSK	737

OY	589	-KRTSEOGVLSSNTRSEOKKRMYYGGHSLDEEDLEMSEPOIKDSVDTCSSITTLNENHS	647
Db	738	VLLPAPQAGRAAPAOBEKKRRR-----ESELQOE-----EERA	772
OY	648	H-----SDKHVTPWQPSKDDGLIGRILLNKLRLKDGSPRDSGAMGLKVVGGKMTESG	701
Db	773	HNOQOEEDROPTWQMOAEKSESRORLSAR-----PLRRQOREQOLAAEENQOEOR	827
OY	702	RLCAFTTKVKKSGSLADTVGHLRPGDEVLEWNGHLLQGAFFEEYNNILLESKPEBOVELV	761
Db	828	FL-----PEEEKEORGR-----ORREKELEOF	851
OY	762	SRPIDIPRIPOSTHAQLESSSSSFOSKMDRPSIVTSPMSPGMLDVPQFLSGOLSIK	821
Db	852	-----LEB-----EBOLRRERPAOOLQEEBDGLODOERROEQORQD	889
OY	822	LMFDKVGHOLLITLGAKDLPRFEDRPRPNRYKIFLPRSDSKNKRRTKVTKTLPEPK	881
Db	890	KM-----RMQL-----EEEKRRRHH-LVAKPALQOLRLKEQQLDOEEBELQ	931
OY	862	NOFTIYSPVNRFEFRERMLETLMQDARVBESEBFLGEILLIETALLDDEPHWY----	937
Db	932	REE-----REKRRROEORQYR-----EEBOLQOEEOQLRE-----BREKRRROERQYRKDK	981
OY	938	KLOTHDVSLLPRPSPYLRPRLOHESP--TRLOKSRKISSEVSDYDCEBGVGAVSD	995
Db	982	KLOQKE-----EOLLGEERPKRRROERKKYREE-----	1011
OY	996	YHNGRDLQSLTVPEQVWSSNHCSPPSGSPHREVDTIGTRTWS--PSAPORPQVNEOG	1052
Db	1012	-----ELQOE-----EOQLREERER-----RRQEWEROYRKKDDELQOEEOQL	1050
OY	1053	HNGTRATGNYNTISRDNRHVMDNHSSDRDRCSEADRPYHRSKSTEQRPILLETTTR	1112
Db	1051	LEER-----EKRRLOERQYREEBELQOEEOQLLEERTTRROELEORYK	1099
OY	1113	SSSEPRDTNLMRSMESLMTGRCAPSPALSRSHPRGVSQUTSPSPSTGRCGRLOPOL	1172
Db	1100	EBELOOEEOQLRE-----EPERKRRROERQORREEBELQOEEOQLLREERK	1147
OY	1173	PKGTLEERSAMDIEERNROKMLNKYQOVAGSDPRLEBDYHSKYRSGMDPH-----	1222
Db	1148	RRROELEORYREEBELOQRKQRYR-----DEDQSDKLQWQEPERKENAVRDNKY	1198
OY	1223	--RGADVTGSKS--SDSDVSOVASVSTSSASRPSSTSYMSVQSEBPRPGNKRKISVFSKM	1279
Db	1199	YKGRNEOFRQLEDSQVRD-----ROSQDLOHLHLEGOERDR-----EOE	1240
OY	1280	NROMVSGKNLTKSTISISGDMCSLEKNDGSQSDTAGALQTSCKKRRSSIGAKKVAIVGL	1359
Db	1241	RRRMOQANHHFPEE-----OLEREKEKA-----KRR-----	1266
OY	1340	SKRSRASOLSQTEGGGKILSTIVOR	1365
Db	1269	DRKSQOEKQLREEREKKRRROETDR	1294

RESULT	3		
RP3A_RAT			
ID	RP3A_RAT	STANDARD:	PRT: 684 AA.
AC	P47709;		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	RABPHILIN-3A.		
GN	RPH3A.		
OS	Rattus norvegicus (Rat).		
OC	Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
OX	NCB1_taxonomy:10116;		
RN	(1)		
RP	SEQUENCE FROM N.A.		

```

RX MEDLINE-95033210; PubMed-7946335;
RA Li C., Takei K., Geppert M., Daniel L., Stenius K., Chapman E.R.,
RA Jahn R., de Camilli P., Suedhof T.C.;
RT "Synaptic targeting of rabphilin-3A, a synaptic vesicle
RT Ca2+/phospholipid-binding protein, depends on rab3A/3C.";
RL Neuron 13:885-898(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 44-167 IN COMPLEX WITH RAB3A.
RC TISSUE=Brain;
RX MEDLINE-99148269; PubMed-10025402;
RA Ostermeier C., Brunger A.T.;
RT "Structural basis of Rab effector specificity: crystal structure of
RT the small G protein Rab3a complexed with the effector domain of
RT rabphilin-3A.";
RL Cell 96:363-374(1999).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED WITH RAS-RELATED
CC PROTEIN RAB-3A IN SYNAPTIC VESICLE TRAFFIC AND/OR SYNAPTIC
CC VESICLE FUSION. COULD PLAY A ROLE IN NEUROTRANSMITTER RELEASE BY
CC REGULATING MEMBRANE FLOW IN THE NERVE TERMINAL.
CC -1- SUBUNIT: MONOMER. SPECIFICALLY EXPRESSED IN BRAIN.
CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: U12571; AAA62662.1;
DR PDB: 1ZBD; 12-APR-89.
DR InterPro: IPR000008;
DR InterPro: IPR001565;
DR InterPro: IPR001965;
DR Pfam: PF00168; C2; 2.
DR Pfam: PF00628; C2; 1.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGM.
DR PROSITE: PS00499; C2_DOMAIN_1; 2.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
DR Repeat: Synapse; protein transport; 3D-structure.
KW DOMAIN 397 645 PHOSPHOLIPID BINDING (PROBABLE).
FT DOMAIN 397 488 C2 DOMAIN.
FT DOMAIN 537 645 C2 DOMAIN.
FT SEQUENCE 684 AA; 75832 MW; 05838BC3C7A86444 CRC64;
SQ
Query Match 3.2%; Score 262.5; DB 1; Length 684;
Best Local Similarity 18.8%; Pred. No. 1.2e-05;
Matches 177; Conservative 105; Mismatches 299; Indels 363; Gaps 38;

```

```

QY 332 NY-----KDSNRGRHSKEYIVDEDEVSDEYERO--RREEYQARY 373
DQ :::::
DB 225 SHGPPTRRASEARMSTTTPDSGWDHGHG-----GGAGPTRSRSPGGEGGLRANSAQAS- 278
QY 374 RSDPNLIARYVVKPQP-----YEDMRTHAEVSRARHREHSHVSIANA 416
DB 279 RPAPASMPSPAPQPQPQPPGSGRAAPGPGRPPEOST-----EAPSPDPGYPGA 328
QY 417 ELEDSTISILNRDPRSGSVSERRAAMENQRSYMERPEAOGQSSYPDRTSNHSPPTP 476
DB 329 V-----APARE-----ERIGPTGFGQAAPITACPIQGAAP 358
QY 477 RSPPIIDPRDMKRADSLRKOHLLDPSAVRKTREKMETMLRNDLSSD-QGSEVSRPP 535
DB 359 AKOPPAEEEEE-----EANSYSDQA--TTGLALEFSLLYDDQNSMLQCTIIRAKG 408
QY 536 PRPHSKKGGKMRQVLSSESEELASTPEYTSDDVLESESYSEKDSQKGRKTEGQ 595
DB 409 LKPMDSN-----GLADPYVKLHLF-----GASKSNKLRTK-- 439
QY 596 VLSDSMTRESEROKKRMYYGGHSLSEEDLEWSEPOIKDSGVDTCSSTTLNEHSHSDKHPT 655
DB 440 --TLRNTRNPNWNETLOYHGIT-EDM-----QRTIIRISVCD-EDKFGHNE----- 482
QY 656 WQPSKDDRLRIGRLILNKLKSDSVPRDSGAMGLKVGSKMTESGRICAPITKVKKGS 715
DB 483 -----FIGE-----TRFSKLKLANQRKRFNIC--LERVIPMKR 514
QY 716 ADTVGHLRDEDEVLENNGRLLQCATFEYEVNIILESPEFOVELVSRPIGDIPIRDST 775
DB 515 AGTTGSGAR-----GMAIYE-----EDQVE-----RIGDIE----- 540
QY 776 HAQLESSSSSFESQKMDRPSISVTPSPMGLRDVPOFLSGOLSTIKLMPKVGHQLIVTI 835
DB 541 -----GKILVSLMYSTQCGGLIYGI 561
QY 836 LGAKDLPSREDGRPRNPVYKIYFLPDRSDKNKRTKTKVTKLEPKNMOTFIYSVHRRF 895
DB 562 IRCVHLAAMDANGSDPFLVLMKLPDMGKRAKAKTKTOIKTKTLNDFNEEFYDIKH-SDL 620
QY 896 REMMLETTIMDOARVRESESPFGEILIELEVALLDDEPRHYK 938
DB 621 AKRSLDISVMDYD--IGKSNDYIGGCOLGI-SAKGERLKHMYE 660
RESULT 4
ID RP3A_BOVIN STANDARD; PRT; 704 AA.
AC 006846;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RABPHILIN-3A.
GN RPH3A.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovine; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.A. AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE-93204952; PubMed-8384302;
RA Shirataki H., Kaibuchi K., Sakoda T., Kishida S., Yamaguchi T.,
RA Wada K., Miyazaki M., Takai Y.;
RT "Rabphilin-3A, a putative target protein for smg p25A/rab3A p25 small
RT GTP-binding protein related to synaptotagmin.";
RL Mol. Cell. Biol. 13:2061-2068(1993).
RN [2]
RP DOMAINS.
RX MEDLINE-94086530; PubMed-8262955;
RA Yamaguchi T., Shirataki H., Kishida S., Miyazaki M., Nishikawa J.,
RA Wada K., Numata S.-I., Kaibuchi K., Takai Y.;

```

```

RT      "Two functionally different domains of rabphilin-3A, Rab3a p25/smg
RT      p25a-binding and phospholipid- and Ca(2+)-binding domains."
RT      J. Biol. Chem. 268:27164-27170(1993).
CC      -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED WITH RAS-RELATED
CC      PROTEIN RAB-3A IN SYNAPTIC VESICLE TRAFFIC AND/OR SYNAPTIC
CC      VESICLE FUSION. COULD PLAY A ROLE IN NEUROTRANSMITTER RELEASE BY
CC      REGULATING MEMBRANE FLOW IN THE NERVE TERMINAL.
CC      -1- SUBUNIT: MONOMER.
CC      -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN.
CC      -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC      CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL collaboration
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      CC
CC      EMBL: D13613; BAA02780.1; -.
CC      DR      HSSP: P21707; IRSY.
CC      DR      InterPro: IPR000008; -.
CC      DR      InterPro: IPR001565; -.
CC      DR      InterPro: IPR001965; -.
CC      DR      Pfam: PF00168; C2; 2.
CC      DR      Pfam: PF00628; PHD; 1.
CC      DR      PRINTS: PR00360; C2DOMAIN.
CC      DR      PRINTS: PR00399; SYNAPTOTAGMN.
CC      DR      PROSITE: PS00499; C2_DOMAIN_1; 2.
CC      DR      PROSITE: PS50004; C2_DOMAIN_2; 2.
CC      KW      Repeat: Synapse; Protein transport.
CC      FT      DOMAIN 1 280 RAB p25/smg p25a BINDING (GTP GAMMA-S-
CC      FT      DOMAIN 281 704 PHOSPHOLIPID BINDING; CA2+-DEPENDENT.
CC      FT      DOMAIN 417 508 C2 DOMAIN:
CC      FT      DOMAIN 577 665 C2 DOMAIN.
CC      SQ      SEQUENCE 704 AA; 77977 MW; 1324D048F58F8FFD4 CRC64;

```

```

Query Match 3.28; Score 261.5; DB 1: Length 704;
Best Local Similarity 18.3%; Pred. No. 1.3e-05;
Matches 172; Conservative 112; Mismatches 307; Indels 351; Gaps 38;

```

```

QY      45 DROKKEEKESVLAKEENKAPRTQWPFSGITELVNVNLOPQOKPKEKEPQTLHQ 104
DB      42 DROKKEE-----LTDEEK-----ETINKVIAAEKME----- 69
QY      105 FEMTKKEQVKKMGESQOOQEO-KGDAPT-CGICHT--KFADGCHNCISYCTKFCARG 160
DB      70 -EMOERGRGLVDLENKRNKRVAGGVNRCILCGFOLGMLGSAC-VVCEDCRKNVCTKCG 127
QY      161 GRVSLRSNKVMVGNLCRKQOEILTKSGAMFYNSGNTLQOPQKVPRLNDEAPQEK 220
DB      128 VETSNNRHPHVLKICIEQREVWKRSGAMFEKG-----PFKQVLPQ-----PMPIK 174
QY      221 AKLHEQPOFGAPDLSPAY--EKGRAHGLTROPTIKNGSEVKHQLASDMSDRK 275
DB      175 KKKPOQPVSEVPRA-APPATPEPKHPARAPTRGDT-EDRRPGQGTGPDMSAERGSGY 232
QY      276 SPVSPODNRRYESESEEDYSQVYPSDGTMPRSPDADRSQREPOPEYEPGHLYND 335
DB      233 GPVYRRASPEAR--HSSGRDSDSMGCHGMAAGDS-----QSPAGL----- 272
QY      336 SNRGHRHSKEYIVDEDEYERKQREERYQARYRSDPNLARYPVKPOPEEOMRI 395
DB      273 -----RRANSVOAS--RPAPASQSPAPDP----- 296
QY      396 HAVYRAHRRHSDVSLANLEDSRISLMDRPSRQSRVSERAAAMENQSYSMERT 455
DB      297 ----- 296
QY      456 READQGSYPORTSNHSPPTPRSPITPLDRPDMRADSLRKOHNLDPSSAVRKTREKME 515
DB      456 ----- 296

```

```

DB      297 -----GQPGPGGSRSPSPGPTGR---FPDQREVPASDP-----DYTGAAAPREERTG 342
QY      516 TMLRNSLSLSDQSESVRP-----PPRPHKSKGGMKQVYSLSS 555
DB      343 GIGGYSAAGTREDRAGPPPSYSTQASAAAPVYASARQPPPEDEBEAN-----STSD 398
QY      556 EEELASTPEYT-----SCDDVESESEYSEKG---DSO-----KGRKRTSEGV 596
DB      399 EATTLALERSLLYDDQNSSLHCTIIKAKGLPMDNSGLADPYKHLILPPASNSNKRT 458
QY      597 LSDSNRSEKRRMYGGHSLDEEDLEWSEBQIKDQSGVDTCSSTLLNEHSHSDKHPTW 656
DB      459 KTLRNTRNPNMNETLVYHGIT-DEDM-----QKTLRLISYVD-----EDKFGHNE----- 502
QY      657 QPSKDDRLIGRLLNKLRLKLDGSPRDSGMLGLKGVGKMTESGRICAFITTKYKKSGLA 716
DB      503 -----FTGE-----TRESLKKLKPNOCKMKNFC--LERYIPMKRA 535
QY      717 DTVGHLPGEVLEWMNGRLLOGATFEFVYNNILLESKEPQVELVYSRPICDIPRIDSTH 776
DB      536 GTTGSAR-----GMALYE-----EEGYE-----RIGDIE----- 560
QY      777 AOLESSSSFEQKMDRPSISVTSPMSGMLRDVPQPLSGQLSIKLMFDKYGHOILYTL 836
DB      561 -----RGKILVIMYSTQOGGLIVGII 582
QY      837 GAKDLPSREGRPNRPVYKIYFLPDRSDKNKRTKYKTLTEPKMNQFITSVPHRRER 896
DB      583 RCYHLAMDANGYSDPVPKLMKLPDMGRKAKHKYQIKRKLTPENEEFEFYDIKH-SDLA 641
QY      897 ERMLEITLMDQARYEESESEFLGLEILELETALDDPEHWYK 938
DB      642 KKSLLDISVWDY--IGKSNDYIGGCQLGI-SAKGERLKHWIE 680

```

```

RESULT 5
APC_HUMAN STANDARD; PRT; 2843 AA.
ID APC_HUMAN
AC P25054; Q15162; Q15163;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ADENOMATOUS POLYPOSIS COLI PROTEIN (APC PROTEIN).
GN APC OR DP2.5.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91335210; PubMed=1651562;
RA Kinzler K.W., Nizard M.C., Su L.-K., Vogelstein B., Bryan T.M.,
RA Levy D.B., Smith K.J., Preisinger A.C., Hedge P., McCreich D.,
RA Fluntz R., Markham A., Groffen J., Boguski M.S., Altshul S.F.,
RA Horii A., Ando H., Miyoshi Y., Miki Y., Nishino I., Nakamura Y.;
RT "Identification of FAP locus genes from chromosome 5q21.";
RL Science 253:661-665(1991).
RN [2]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RC TISSUE=Petal brain;
RX MEDLINE=91330307; PubMed=1678319;
RA Joslyn G., Carlson M., Thliveris A., Albertsen H., Gelbert L.,
RA Samowitz W., Groden J., Stevens J., Spirio L., Robertson M.,
RA Sargent L., Kravcho K., Wolff E., Burt R., Hughes J.P.,
RA Warrington J., McPherson J.D., Wasmuth J., Le Paslier D.,
RA Abderrahim H., Cohen D., Leppert M., White R.;
RT "Identification of deletion mutations and three new genes at the
RT familial polyposis locus.";
RL Cell 66:601-613(1991).
RN [3]
RP ASSOCIATION WITH CATENINS.
RX MEDLINE=94082295; PubMed=8259519;
RA Su L.-K., Vogelstein B., Kinzler K.W.;

```

RT "Association of the APC tumor suppressor protein with catenins.";
 RL Science 262:1734-1737(1993).
 RN [14]
 RP REVIEW ON VARIANTS.
 RX MEDLINE-94134728; PubMed-8111410;
 RA Nagase H., Nakamura Y.;
 RT "Mutations of the APC (adenomatous polyposis coli) gene.";
 RL Hum. Mutat. 2:425-434(1993).
 RN [15]
 RP VARIANTS FAP.
 RX MEDLINE-91335211; PubMed-1651563;
 RA Nishisho I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A.,
 RA Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J.,
 RA Petersen G., Hamilton S.R., Nilbert M.C., Levy D.B., Bryan T.M.,
 RA Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.;
 RT "Mutations of chromosome 5q21 genes in FAP and colorectal cancer
 RT patients.";
 RL Science 253:665-669(1991).
 RN [16]
 RP VARIANTS FAP.
 RX MEDLINE-93265030; PubMed-1338904;
 RA Miyoshi Y., Nagase H., Ando H., Ichii S., Nakatsuru S., Aoki T.,
 RA Miki Y., Mori T., Nakamura Y.;
 RT "Somatic mutations of the APC gene in colorectal tumors: mutation
 RT cluster region in the APC gene.";
 RL Hum. Mol. Genet. 1:229-233(1992).
 RN [17]
 RP VARIANTS FAP.
 RX MEDLINE-93244793; PubMed-1338691;
 RA Nakatsuru S., Yanagisawa A., Ichii S., Tahara E., Kato Y.,
 RA Nakamura Y., Horii A.;
 RT "Somatic mutation of the APC gene in gastric cancer: frequent
 RT mutations in very well differentiated adenocarcinoma and signet-ring
 RT cell carcinoma.";
 RL Hum. Mol. Genet. 1:559-563(1992).
 RN [18]
 RP VARIANTS FAP W-1348, AND VARIANTS D-1118; M-1292; V-1304 & S-2502.
 RX MEDLINE-93250848; PubMed-1338764;
 RA Nagase H., Miyoshi Y., Horii A., Aoki T., Petersen G.M.,
 RA Vogelstein B., Maher E., Ogawa M., Maruyama M., Utsunomiya J.,
 RA Baba S., Nakamura Y.;
 RT "Screening for germ-line mutations in familial adenomatous polyposis
 RT patients: 61 new patients and a summary of 150 unrelated patients.";
 RL Hum. Mutat. 1:467-473(1992).
 RN [19]
 RP VARIANTS FAP TRP-99.
 RC TISSUE-Peripheral blood lymphocytes;
 RX MEDLINE-95134544; PubMed-7833149;
 RA Dobbie Z., Spycher M., Huerliman R., Ammann R., Ammann T., Roth J.,
 RA Mueller A., Mueller H., Scott R.J.;
 RT "Mutational analysis of the first 14 exons of the adenomatous
 RT polyposis coli (APC) gene.";
 RL Eur. J. Cancer 30A:1709-1713(1994).
 RN [10]
 RP VARIANTS FAP GLY-722.
 RX MEDLINE-95135430; PubMed-7833931;
 RA Stella A., Montero M., Resta N., Marchese C., Susca F., Gentile M.,
 RA Romio L., Pilla S., Prete F., Mareni C., Guanti G.;
 RT "Four novel mutations of the APC (adenomatous polyposis coli) gene in
 RT FAP patients.";
 RL Hum. Mol. Genet. 3:1687-1688(1994).
 RN [11]
 RP ERRATUM.
 RA Stella A., Montero M., Resta N., Marchese C., Susca F., Gentile M.,
 RA Romio L., Pilla S., Prete F., Mareni C., Guanti G.;
 RL Hum. Mol. Genet. 3:1918-1918(1994).
 RN [12]
 RP VARIANTS FAP ILE-171.
 RX MEDLINE-97144176; PubMed-8990002;
 RA van der Luijt R.B., Meera Khan P., Vasen H.F.A., Tops C.M.J.,
 RA van Leeuwen-Correllisse I.S.J., Wijnen J.T., van der Klift H.M.,
 RA Plug R.J., Griffioen G., Fodde R.;
 RT "Molecular analysis of the APC gene in 105 Dutch kindreds with

RT familial adenomatous polyposis: 67 germline mutations identified by
 RT DGGE, PTT, and southern analysis.";
 RL Hum. Mutat. 9:7-16(1997).
 RN [13]
 RP VARIANTS LYS-1307.
 RX MEDLINE-98400248; PubMed-9731522;
 RA Redston M., Nathanson K.L., Yuan Z.O., Neuhausen S.L., Satagopan J.,
 RA Wong N., Yang D., Nafa D., Abrahamson J., Ozcelik H.,
 RA Antin-Ozerkis D., Andrulis I., Daly M., Plinsky L., Schrag D.,
 RA Gallinger S., Kaback M., King M.-C., Woodage T., Brody L.C.,
 RA Godwin A., Warner E., Weber B., Foulkes W., Offit K.;
 RT "The APC 11307K allele and breast cancer risk.";
 RL Nat. Genet. 20:13-14(1998).
 RN [14]
 RP VARIANTS LYS-1307 AND GLN-1317.
 RC TISSUE-Peripheral blood;
 RX MEDLINE-98393712; PubMed-9724771;
 RA Frayling I.M., Beck N.E., Ilyas M., Dove-Edwin I., Goodman P.,
 RA Pack K., Bell J.A., Williams C.B., Hodgson S.V., Thomas H.J.W.,
 RA Talbot I.C., Bodmer W.F., Tomlinson I.P.M.;
 RT "The APC variants 11307K and E1317Q are associated with colorectal
 RT tumors, but not always with a family history.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10722-10727(1998).
 RN [15]
 RP VARIANTS LYS-1307.
 RX MEDLINE-98400259; PubMed-9731533;
 RA Woodage T., King S.M., Wacholder S., Hartge P., Struwing J.P.,
 RA McAdams M., Laken S.J., Tucker M.A., Brody L.C.;
 RT "The APC 11307K allele and cancer risk in a community-based study of
 RT Ashkenazi Jews.";
 RL Nat. Genet. 20:62-65(1998).
 RN [16]
 RP VARIANTS LYS-1307.
 RX MEDLINE-99138651; PubMed-9973276;
 RA Gryfe R., Di Nicola N., Lai G., Gallinger S., Redston M.;
 RT "Inherited colorectal polyposis and cancer risk of the APC 11307K
 RT polymorphism.";
 RL Am. J. Hum. Genet. 64:378-384(1999).
 RN [17]
 RP VARIANTS FAP CYS-1171 & THR-2738, AND VARIANTS GLY-1057 & VAL-1822.
 RX MEDLINE-9913859; PubMed-9950360;
 RA Wallis Y.L., Morton D.G., McKeown C.M., MacDonald F.;
 RT "Molecular analysis of the APC gene in 205 families: extended
 RT genotype-phenotype correlations in FAP and evidence for the role of
 RT APC amino acid changes in colorectal cancer predisposition.";
 RL J. Med. Genet. 36:14-20(1999).
 RN [18]
 RP VARIANTS FAP PRO-1184.
 RX MEDLINE-99401091; PubMed-10470088;
 RA Iamlium H., Ilyas M., Rowan A., Clark S., Johnson V., Bell J.A.,
 RA Frayling I.M., Efstathiou J., Pack K., Payne S., Roylance R.,
 RA Gorman P., Sheer D., Neale K., Phillips R., Talbot I.C., Bodmer W.F.,
 RA Tomlinson I.P.M.;
 RT "The type of somatic mutation at APC in familial adenomatous polyposis
 RT is determined by the site of the germline mutation: a new facet to
 RT Knudson's 'two-hit' hypothesis.";
 RL Nat. Med. 5:1071-1075(1999).
 RN [19]
 RP FUNCTION: TUMOR SUPPRESSOR. ALLOWS THE RAPID TURNOVER OF BETA-
 CATENIN. APC ACTIVITY CORRELATED WITH ITS PHOSPHORYLATION STATE
 CC ALLOWS THE DOWNREGULATION OF CYTOPLASMIC BETA-CATENIN.
 CC SUBUNIT: FORMS HOMOLOGOMERS AND ASSOCIATES WITH CATENINS.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES TYPES.
 CC -1- PTM: PHOSPHORYLATED BY GSK-3B.
 CC DISEASE: DEFECTS IN APC ARE A CAUSE OF FAMILIAL ADENOMATOUS
 CC POLYPOSIS (FAP) AND GARDNER'S SYNDROME (GS), THAT CONTRIBUTE TO
 CC TUMOR DEVELOPMENT IN PATIENTS WITH NONINHERITED FORMS OF
 CC COLORECTAL CANCER. FAP IS CHARACTERIZED BY ADENOMATOUS POLYPS OF
 CC THE COLON AND RECTUM, BUT ALSO OF UPPER GASTROINTESTINAL TRACT
 CC (AMPULLARY, DUODENAL, AND GASTRIC ADENOMAS). THIS IS A VICIOUSLY
 CC PREMALIGNANT DISEASE WITH ONE OR MORE POLYPS PROGRESSING THROUGH
 CC DYSPLASIA TO MALIGNANCY IN UNTREATED GENE CARRIERS WITH A MEDIAN


```

CC      AGCAT DIAGNOSIS OF 40 YEARS.
CC      -I DISEASE: APC. MUTATIONS HAVE LED TO SOME INTERESTING OBSERVATIONS.
CC      (1) THE GREAT MAJORITY OF THE MUTATIONS FOUND TO DATE WOULD RESULT
CC      IN TRUNCATION OF THE APC PRODUCT. (2) ALMOST ALL THE MUTATIONS
CC      HAVE OCCURRED WITHIN THE FIRST HALF OF THE CODING SEQUENCE, AND
CC      SOMATIC MUTATIONS IN COLORECTAL TUMORS ARE FURTHER CLUSTERED IN A
CC      PARTICULAR REGION, CALLED MCR (MUTATION CLUSTER REGION). (3) MOST
CC      IDENTIFIED POINT MUTATIONS IN THE APC GENE ARE TRANSITIONS FROM
CC      CTG/STG TO OTHER NUCLEOTIDES. (4) THE LOCATION OF GERM-LINE
CC      MUTATIONS TENDS TO CORRELATE WITH THE NUMBER OF COLORECTAL POLYPS
CC      IN FAP PATIENTS. INACTIVATION OF BOTH ALLELES OF THE APC GENE
CC      SEEMS TO BE REQUIRED AS AN EARLY EVENT TO DEVELOP MOST ADENOMAS
CC      AND CARCINOMAS IN THE COLON AND RECTUM AS WELL AS SOME OF THOSE IN
CC      THE STOMACH.
CC      -I SIMILARITY: CONTAINS 7 ARM REPEATS.
CC      -I DATABASE: NAME=APC; NOTE=Information about APC mutations;
CC      WWW="http://perso.cuie.fr/thierry.sousai/apc.html".
CC      -----
CC      Query Match      3.1%, Score 258.5, DB 1; Length 2843;
CC      Best Local Similarity 17.9%; Pred. No. 0.00011;
CC      Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69.
CC
QY      30 SHLTDEERKILAVMDRCKEERKQSVLAKIEKHAQDTQMPFSGITELVNNVLPQO 89
DB      1012 NMMDNDGELDTPIVSLKYSDQ-----LNSGRSPSONEMRARRKHILE--DEIKQSD 1065
QY      90 KQPKNEKQETKLHQOFEMXKEQVKKMGESQO---QOEKKGAPTCGICIKTKFADGCH 146
DB      1066 ROSRNS-----TTPYPTTESTDKHLKPFHNFQOE-----CVSYRRSGANGSET 1112
QY      147 NCSYCTQFCARCGGRVSLRSKNVAVNCMLCKRQOEILTKSCAMFYNSGNTLQOPQY 206
DB      1113 N-----RVGNHNGINQVQSOLC-----QEDDYEDDKPTNSERYSEEEQHE-- 1155
QY      207 PGLANEAPQEKKAKLHEQPOFGARGLSVPAYEKGAGHGLTRQDTLKNSGVKHOIA 266
DB      1156 -----EEKPTNYSTIKYNDKRNHVDPIDYSL-----KTA 1184
QY      267 SDMPSDRKSPSVSHQDNRRYEOSEREDYSOYVSDGTMPRSPSDYARRSQREPFYE 326
DB      1185 TDIPSSQKQSFSEKSSSSQSSKTE-----HMSSSNTSTSPSSNAKRQNLHSSMQ 1237
QY      327 EFGHLNRYDRSNRGRHRSKE---YIVD-----EDVE 355
DB      1238 SRSGOPKAKATCKVSSINQETIQTCVEEDTPICFSCSSLSLSSADEIGCNOQTQEAD 1297
QY      356 SRDEYERQREDEYQARYSDPNLARYPKRPQPYEQOMTHAEVSNARHERHSDVSLAN 415
DB      1298 SANTIQLAIEIKKICTRSAEDVSEVPAYSOHPRKSSRLQGS-SLSSSARKAAVEFSS 1356
QY      416 AELEDSRISILMDRP-----SRQSVSERRAMENQRSYMERGTREAO----- 459
DB      1357 GAKSPKSAQATPKRPPEHYVETPLMFRCSTVS---SLDSFESRSTASSVQSEPCSG 1412
QY      460 -----GGSYPQRTSNHSPTPRRSPIPLDRPDMRRADSLRKQNLHLDSSAVAKTKREK 513
DB      1413 WVGSIISPDLPDPSQGTWMPRSPRSKTRPP-----PQTA--QTKRE- 1451
QY      514 METMLRNDLSLSDQSESVAPPPRPRAKSKGGMROYSLSSSEEL-----ASTPEYTS 568
DB      1452 ---VKNKAPLTAEKRES-----GPKQAAVNVAOVQVLPADLTLLHETESTPGGFG 1502
QY      569 D-----DVEL-----ESESVEKGDGO-GKRT--SEOG 595
DB      1503 SSSLSALSLDEPFQKDVLRIMPPYQEDNDNGNETSEDPKESNENQKEAEKTIIDSEKD 1562
QY      596 VLSDSN-----TRSERQKRMYYGGHSLDEEDLEWSEPOIKDSGVDTGSS 639
DB      1563 LLDSDSDDDDEITILECIIISAMPTKSSRAKKAKRAQTSKLPPVARKPSQL-----PYVK 1616
QY      640 TLLNEHSHSKAHPTVWQSKDGRLL--IGRIILN-----KRLKDGSVPRDSGMLGL 690

```

Db	1617	LLSQNRLOPKH-VSFTFGDDMPRVYCEGPIINRSTASTLSLDLIESPPNLAAGEV	1675
Qy	691	KVVGKMTESGRCLAFITVYKKGSLADYGHILRPGDEVLEWNRGLLOGATFEVYNTILE	750
Db	1676	R--GG--AOSG-----EFEXKDTIPTEG--RSTDEA-----OGGTSSVTIPELD	1714
Qy	751	SKPEQVELVY-----SRIGD-----IPRIDSHAOLESSSSSESKMDRPSISVT	799
Db	1715	DNKAEGDGLIACINSAMPRGSHKPFRRYKILMDOVO-QASASSAPAKNOLOGKKKKPT	1773
Qy	800	SPSPGMLNDVPOFLSGQSLIKLMPDKVGHOLIIVTILGAKDLPSREDGRPNRYVKIYFL	859
Db	1774	SPVKP-----IQNTFYRRRV-----RKNADESKNNMARVF	1805
Qy	860	PDRSDKNRRRTYVKKTLBPK-----WNQTIYSVNH-----RR	893
Db	1806	SNKDSKKONLKNNSKDPNDKLPNNEDRVGSPAFSPHHYITIEGPHYCFSRNDSLSL	1865
Qy	894	ERRERLETTMDQAVNREESEFLEIILETALLDDEPHMYKLIOTHVSSILPLR--	951
Db	1866	DDDDDDVDLSR-EKAEILRAKENESEAKTSTHETLSNOQSANIKQYA-IARQPINRGQ	1922
Qy	952	PSPYLPEROL-----HGESPIRLOR-----SKRISDESDVDC-----	987
Db	1923	KPFILOKOSTFPOSSKDIDPRGATDEKILQNFALIENTPYCFSHNSSLSLSDIDQENNK	1982
Qy	988	-----DGVVSVDRHNG-----RDLQSTLVSPEQYSSNHC	1020
Db	1983	ENBPICKETEPPOQGBPRKQASGVAPKSFHVEDTYPGCFRNSLSLSLSDIEDLLQEC	2042
Qy	1021	SPSGSHRVD---VIRTSWSAPSAPPOANVOGHRGTATGHYNTISM-D-RHRMDD	1076
Db	2043	ISSAMKRRKKPSRLKGDNEKHP-----RNM-GGILGEDLTLDKIDQRPDSEHGSPD	2095
Qy	1077	HXSDDRCDEAD-----RQ-----PYHRSKTE	1101
Db	2096	SENEFDKALIOEGNLSIVSSLHQAAAACLSRQASSDSDSLILSKGSLGSPFHLTPDOE	2155
Qy	1102	QRFLL-----ERTTNRSSSERPDTN-----	1122
Db	2156	EKPFTSNKGPRIKLPEKSTLETLETKKIESESKIGKKVYKSLITGKVRNSEISGQMKQ	2215
Qy	1123	LMRSPSLMTGSAAPPSPALNSHRTGY-----QTSPTPTGRRGROLPLPLP	1174
Db	2216	PLQANNPISIRGRTMHIDPVRNSSSTSPSKKGPLKTPAKSPSEG---QTATTPS	2271
Qy	1175	KGLERSAMDIEERNOMKLNKKOYVAGSPRLQDVHNSYRSQ--WDPHRGAD-----	1226
Db	2272	RGAKPVPKSELSPVAKOT-----SQJGGS-----KAPSRGSRDSTPSRPQOPLSR	2319
Qy	1227	TVSTKSSDSDVDSAVASRTSSARFSTSYMSYOSERPGRNRIYFTSKMQNRQGV	1286
Db	2320	PLQSPERNSTISPERNISTSPNKLSQLPRISPSPTASTKSSGSGMS-YTS--PERQM--S	2374
Qy	1287	GKNLTKSTISGDMCLEKNDGQSODTAVGALGTSGKKRRSSIGAKKVAIYGLSKRBSA	1346
Db	2375	QONLTKOTGLSKMASSIPRSESAASKGLNOMNONGANKK-----VELSRMSTK	2423
Qy	1347	SQLSQTEGGGKKL-----RSTVGRSTETGLAVEMRNMNTQASRST-----	1388
Db	2424	SSGSESDRESEPLVKNQSTFIKAPSPFLLRKKLLEESASFESLSPSRAPSTRSOAOPTV	2483
Qy	1389	-----DGSNYSYS	1397
Db	2484	LSPSLPLDMLSTHSS	2498
RESULT	6		
TPR_HUMAN	ID	TPR_HUMAN	STANDARD; PRT: 2349 AA.
AC	P12370:		
DT	01-OCT-1989 (rel. 12, Created)		
DT	01-OCT-1996 (rel. 34, Last sequence update)		

01-OCF-2000 (Rel. 40, Last annotation update)
 DE NUCLEOPROTEIN TPR.
 GN TPR.
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID:9606;
 RN [1]
 RN MEDLINE=93064711; PubMed=1437155;
 RX MEDLINE=93064711; PubMed=1437155;
 RA "The human tpr gene encodes a protein of 2094 amino acids that has
 RT extensive coiled-coil regions and an acidic C-terminal domain."
 RT Oncogene 7:2329-2333(1992).
 RN [2]
 RN REVISIONS AND CHARACTERIZATION
 RX MEDLINE=95096166; PubMed=7798308;
 RA Byrd D.A., Sweet D.J., Pante N., Konstantinov K.N., Guan T.,
 RA Saphire A.C.S., Mitchell P.J., Cooper C.S., Aebi U., Gerace L.;
 RT "tpr, a large coiled coil protein whose amino terminus is involved in
 RT activation of oncogenic kinases, is localized to the cytoplasmic
 RT surface of the nuclear pore complex."
 RT J. Cell Biol. 127:1515-1526(1994).
 RN [3]
 RN SEQUENCE OF 1-142 FROM N.A.
 RX MEDLINE=88262257; PubMed=3387099;
 RA King H.W.S., Tempest P.R., Merrifield K.R., Rance A.J.;
 RT "tpr homologues activate met and raf".
 RT Oncogene 2:617-619(1988).
 CC -1- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILIBRIS OF THE NUCLEAR PORE
 CC COMPLEX IMPLICATED IN NUCLEAR PROTEIN IMPORT. ITS AMINO TERMINUS
 CC IS INVOLVED IN ACTIVATION OF ONCOGENIC KINASES.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF THE NUCLEAR PORE
 CC COMPLEX. THE ASSEMBLY OF THE NPC IS A STEPWISE PROCESS IN WHICH
 CC TPR-CONTAINING PERIPHERAL STRUCTURES ASSEMBLE AFTER OTHER
 CC COMPONENTS, INCLUDING P62.
 CC -1- TISSUE SPECIFICITY: HIGHEST IN TESTIS, LUNG, THYMUS, SPLEEN AND
 CC BRAIN, LOWER LEVELS IN HEART, LIVER, AND KIDNEY.
 CC -1- DISEASE: INVOLVED IN TUMORIGENIC REARRANGEMENTS WITH THE MET, TRK
 CC OR RAF GENES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X66397; CAA47021.1; -;
 DR EMBL: Y00672; CAA6681.1; -;
 DR PIR: S00928; S00928.
 DR MIM: 189940; -;
 KW Heptad repeat pattern; Coiled coil; Proto-oncogene;
 KW Chromosomal translocation; Nuclear protein; Transport.
 KM
 FT DOMAIN 78 360 COILED COIL (POTENTIAL).
 FT DOMAIN 422 571 COILED COIL (POTENTIAL).
 FT DOMAIN 575 628 COILED COIL (POTENTIAL).
 FT DOMAIN 758 805 COILED COIL (POTENTIAL).
 FT DOMAIN 834 869 COILED COIL (POTENTIAL).
 FT DOMAIN 934 979 COILED COIL (POTENTIAL).
 FT DOMAIN 1004 1064 COILED COIL (POTENTIAL).
 FT DOMAIN 1138 1166 COILED COIL (POTENTIAL).
 FT DOMAIN 1196 1241 COILED COIL (POTENTIAL).
 FT DOMAIN 1262 1304 COILED COIL (POTENTIAL).
 FT DOMAIN 1354 1434 COILED COIL (POTENTIAL).
 FT DOMAIN 1476 1595 COILED COIL (POTENTIAL).
 FT DOMAIN 527 530 POLY-SER.
 FT DOMAIN 1833 1836 POLY-GLU.
 FT DOMAIN 1957 1964 POLY-ASP.
 FT DOMAIN 2295 2298 POLY-SER.
 SQ SEQUENCE 2349 AA; 265600 MW; AFDD6895CEDCAC9EF CRC64;

Query Match 3.18; Score 255.5; DB 1; Length 2349;
 Best Local Similarity 18.18; Pred. No. 0.00012;
 Matches 341; Conservative 268; Mismatches 675; Indels 599; Gaps 76;
 10 RPAPTASOPPPPEMDLSHL-TEERKIIILVMDROKKEEKEQSVLKIEBHKAQP 68
 644 RPSTQIVSTPAPYVPIESTALIEKAKLQDELFEYKKKEKANE----- 690
 69 TQWPFSGITELVNNVLOPQOKPNEKEPQ-TKLHQ-----PEMKKEQKKMGEE- 118
 691 -----KIQNEOLEKLOEYVTDLSQNTKISTQDFASKREEMQDNVEGRREIT 740
 119 SQOOOEQKGDAPTCIGICKTKFADGCGHNCSCYQKFCARCGRVLSSNKVMVYCNLCR 178
 741 SLHERNQKLTATTT-----Q 754
 179 KOQELITKSGAMFYNSGNTLOPQKVP-RGLRNEEAPQEK-----KAKLHEQPFQ 231
 755 KOQELI-----NTMQLRGANEKLAVALVRAENLKEKEEMKLSEVRLSQORE--- 803
 232 APGLISVAVEKGRAGHLTRDTI-----KNGSGVKHOIASDMPDRKSPSVSRDQNR 286
 804 -----SLAEORGQMLLTNLQTLQIGILERSETETKQRLSSQI---EKLHEHSLKKKL 855
 287 YEQSEERE-----DYSQVPSDGTMPSPSDYADRSRQREPPQFEEPGHLNYRD 335
 856 ENEVQRHITLNLNDVQLDTRKQDTEFNLHLNTEKLKNAQKEIATLQKHLSNMEVOY 915
 336 SNRGRHSKEYIVDEDEYSEDEYERQREBEYQARYR---SDPLIARYPV----- 384
 916 ASQSQRTGCKGQPSNKEDVDLVSQRLQTEQVNDLKERIKTSTNVEQIOAMVTSLEES 975
 385 --KQPYEQRMTAEVS-----RAHRRHSQVSLANAELEDSR-----ISL 425
 976 LNKQVTEVRKNIEVRLKESAEFQTOLEKLMVEKEKEQELQDDKRAIESMEQOLSE 1035
 426 LMDPRRSQSVSERRAMENQSYSMERTREAOQSGSYQRTSNHSPPTPRSPPLDRL 485
 1036 LKLTSSVQNEVOEALQRASTALSNQOARDCQEQAKTAVEAQN---KYREMLLHA 1090
 486 PMRRADSLRQ-----HHLDPSAVRKT-----KREKM-----E 515
 1091 ADVEALQAAKEQVSKASVROHLETTQAKESQULECKASWEERERILAKDEVSCVGRCE 1150
 516 TMLRNDLSQDSQSVRPPPRPHKSKGKMRQVSLSSSE----- 557
 1151 DLEKQRLHLDQLEKL---SDKVVASVKEGVQGLNLSLEEGKSQQLLEILRFRREK 1207
 558 -----ELASTPEYISCDVLESESVSEKQDSQKKR-----KTS 592
 1208 EIAETREFEVAQVESLRKYRQVELLERELQLEEDSLNAREKVOYTAQMAOHELMKTE 1267
 593 EGVGLSDSNTRSEROKKRMYVGGHSLSEDELEWSEPOITKDSQVDTCSSTTLNEHSHS--- 649
 1268 TMYVMETNMLRPEKR-----LEQDILQOMQAKYRKLELDILPLQOANAELESEKSGM 1320
 650 -----DKHPVTQPSKQDRLIGRLILNKLKLDGSGVPRDSDGAMGLK 691
 1321 LQAEKKLLEEDVKKRKARNQHLVSOQ---KQPDTEYRLILSEK-----EYHTK 1366
 692 VVGKMTESGRICAFILTKVKKG-----SLADYVGLRPGDEVLEBN----- 732
 1367 RIQQLTEIGRLKAEIKASNSALNNONLQSLKEDLKLKVTETKTIKQDADKIIDQIE 1426
 733 -----GRLLQ-----GATFEVYNIILSEKPEQVELVSPRIGDIPRIPIST 775
 1427 KVKTITQYKKGKRYKQYELKKAQODKVMETSAQSSGDHDEQHVSDQEMDEL---KET 1482
 776 HAQLESSSSSFESQ-----KMDRSISVTSPMSPGMKRDVQFQLSGQ-- 817
 1483 LQAEITKSKLSQOVENLQKTLSEKETEARNLQEQTVQLQSLSELS---RLKQDLQDRTQPE 1540

```

OY 818 -----LSIKIMFDKVGHOVLITLGAKDLPSREDG--RPNRPVYKIFLPD 861
DB 1541 QLRQOITEKEKTKRAIYAASKIAH-----LAGVKDOLTKRENEELKQNG-----ALDQ 1590
OY 862 RSDKNKRTKTVKKTLEPKMNOTFIYSPVHR--REFREMEITLMDQARVEESEFLG 919
DB 1591 QKDELVDVITALKSQEGR-----ISRLELRREHOENHLE-----QRDEPOEPSN 1636
OY 920 EI-----LIEET-----ALLDDEPHWKLOTHVYSLPLRPPSYLRRLOHGRS 965
DB 1637 KVPQOQOITLTKTPPASGERGIATSDPP-----TANKPPVYVSTPSKVTAAAMAGNK 1690
OY 966 PTRRLQSRKRISDEVSUYDEEDGVGVVDYHNGRDLQSLTSLVEQYMNCSHSCSS 1025
DB 1691 STPRASIRPMVTPATVY-----NPTTPPTAIVMPTTYQVESQEAAMSSEEP 1734
OY 1026 PHRYVDYIGRT---RSMSPSAP---PQRYNEQGHRTGRATGHYNTISRMDRHRVMDH 1077
DB 1735 VEHPVFEFGSTGSVASTSPNVQPSISQPLIVQO--QTOAT----- 1773
OY 1078 YSSDRDRCEAADRQPIHRSRSTEDRPLERT-----TTSRSSERPDTNLMRSM 1127
DB 1774 -----AFVQPTQOS-----HPQIEPANOELSNIVEVQSSPYERSTG----- 1812
OY 1128 PSLMTGRSAPPS---PALRSRHPRTGSVQTSSTPGTGR--GRQLPOLPKGTLERSA 1182
DB 1813 TAVGTVSATSSSLPKRTREEDSTIEASDVQDDIYEMLPKTKLVTVPGT--EEEV 1871
OY 1183 MDIEERNOMKLNKYQVAGSDPRLQDYHSKYRSG---WDPHKGADTVSTKSSDSV 1237
DB 1872 MAESTDEVEIYQY-----NDSQDSIGCVTQGDVTPMEDSBETS--QSLQIDL 1920
OY 1238 SDVAVASTSSASRSTSYSVQSEPRGRNKKISVFTSKMNRQ-----MGVS 1286
DB 1921 GPLQSDQDTTTS-----QDQGRKGDVIVIDSDDEDEDEDDDDDDDTGMGE 1970
OY 1287 GKNLTSTKSTISGDMCKLEKNDGOSQDFAVGALGTS--GKKRSGSIGAKVAIVGLSRKRS 1345
DB 1971 GEDSNIEGIG---SAGNDGYEADDAEGGDTGPTTESMG-----GGEQNHRA 2017
OY 1346 ASQSLQTEGGGKKLST---VQSTETGLAVEMRMNMTROASRESTGDSMNSYSEGNL 1401
DB 2018 ADSQSGEGNTGAABSSRSQEVRSQPSASAE-----RQAPR-----APQSPRRPH 2066
OY 1402 IFPGVRLASDSQSFGLGLP-----AQLVGRQTLATPAMGDIVGMKKGQLE 1452
DB 2067 LPPLRLTHAPQ-----ELGPPVQRIQMTROSQVGRGLQTPGIGKMOHFEDDED-- 2118
OY 1453 VEIIRARGLVYKPGSKTLBAPVVKYUULLDNGVCIAKKKTVA---RKTEPLVQQLSFE 1509
DB 2119 -----TVPSTPLVVRH-----RTDGFABAIHSPQVAGVPRFRGPP-----E 2156
OY 1510 ESPQGRVLIITVGDYGHMDKHSFNGVAQ---LLDLELISNMVIGWFKL---FPPSSLY 1563
DB 2157 DMPQYSSSH---SBLQGLASQGGIGLWETPLFLAHHESSGGRSVPTTLQVAAAPVTVFT 2212
OY 1564 DPTSAPLTRRASQS--SLESSTG 1584
DB 2213 ESTTSASEHASQSVPMVTTSTG 2235

```

```

RESULT 7
APC_RAT STANDARD. PRT: 2842 AA.
AC P70478;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ADENOMATOUS POLYPOSIS COLI PROTEIN (APC PROTEIN).
3N APC.
3S Rattus norvegicus (Rat).
3C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.

```

```

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER 344/N; TISSUE=Brain;
RX MEDLINE=96116966; Pubmed=8563176;
RA Toyota M., Ushijima T., Kakiuchi H., Watanabe M., Imai K., Yachi A.,
RT Sugimura T., Nagao M.;
RT "cDNA cloning of the rat APC gene and assignment to chromosome 18.";
RL Mamm. Genome 6:746-748(1995).
RN [2]
RP MUTAGENESIS.
RC STRAIN=SPRAGUE-DAWLEY, AND FISCHER 344/N;
RX MEDLINE=95148647; Pubmed=7846077;
RA Kakiuchi H., Watanabe M., Ushijima T., Toyota M., Imai K.,
RT Weisburger J.H., Sugimura T., Nagao M.;
RT "Specific 5'-GGA-3'->5'-GGA-3' mutation of the APC gene in rat colon
RT tumors induced by 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:910-914(1995).
CC -1- FUNCTION: TUMOR SUPPRESSOR. ALLOWS THE RAPID TURNOVER OF BETA-
CC CATEININ. APC ACTIVITY CORRELATED WITH ITS PHOSPHORYLATION STATE
CC ALLOWS THE DOWNREGULATION OF CYTOPLASMIC BETA-CATEININ (BY
CC SIMILARITY).
CC -1- SUBUNIT: FORMS HOMODIMERS AND ASSOCIATES WITH CATEININS (BY
CC SIMILARITY).
CC -1- PTM: PHOSPHORYLATED BY GSK-3B (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 7 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D38629; BAA07609.1; .
DR HSSP: Q02248; 2BCT.
DR InterPro: IPR000225; .
DR Pfam: PF00514; Armadillo_seg; 4.
DR PROSITE: PS50176; ARM_REPEAT; 1.
KW Anti-oncogene; Phosphorylation; Coiled coil; Repeat.
FT DOMAIN 1 728 LEU-RICH.
FT DOMAIN 1 62 COILED COIL (POTENTIAL).
FT DOMAIN 1 260 COILED COIL (POTENTIAL).
FT REPEAT 451 493 ARM 1.
FT REPEAT 503 545 ARM 2.
FT REPEAT 546 589 ARM 3.
FT REPEAT 590 636 ARM 4.
FT REPEAT 637 681 ARM 5.
FT REPEAT 682 723 ARM 6.
FT REPEAT 724 765 ARM 7.
FT DOMAIN 739 2831 SER-RICH.
FT DOMAIN 1130 1155 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 1556 1575 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 1864 1891 HIGHLY CHARGED.
FT MUTAGEN 523 523 C->R: IN AN TO-INDUCED COLON TUMOR.
SQ SEQUENCE 2842 AA: 310530 MW: 3CB92E8A54EBF47 CRC64;

```

```

Query Match 3.1%; Score 255; DB 1; length 2842;
Best Local Similarity 19.5%; Pred. No. 0.00016;
Matches 333; Conservative 204; Mismatches 636; Indels 532; Gaps 78;
OY 20 PPGQ-----PEMDLSHTEERKIIILAVDRQKKEEKESV-----LK 59
DB 1525 PAVQENDNGNETPEQPEESN-----ENQDEVEKPESEKDLDDSDDDIE 1571
OY 60 IKKE--HKAQTYQMFPPSGIETLVNNTQPOKQENKEPQTKLHQEFMYEYKKNKE 117
DB 1572 ILECIISAMPTK-----SSRAKKILAQATSKLP--PPVAKRPSQLPVY---KLPS 1618
OY 118 ESQOOOEOK-----GDAPTGICHTKTFADCGCHNCYCOT-----KFCARGC 160

```

Db	1619	QSRLOAQKHYSEFTQDDPV	-----RYVCVEBETPIINFSTAIYSLSDLIETSPENNELAAGG	1672
Qy	161	GRVLSRKKVMYVNCILCRQOEILTKSGAMFYNGSGNTLQDPQKVPGRLENEAPQEQK	220	
Db	1673	VRASVOJSEF	-----EKRDITPIEG-----RSTDEAORCK	1702
Qy	221	AKLHEQPOFOGA	---PGLSV---PAYEKGRAHGLTR-----ODPIKNGSGV-KHQ	264
Db	1703	VSSIAIPDLDSKAEEGDILAECLNSALPDKGSHKPFYKATIMDOYQQAQSMNMSGINKNQ	1762	
Qy	265	IASDMPSDRKRSVSRDONRREY-OSEEREDYSOYVPSDGTMPNSPSDYADRDSOREPO	323	
Db	1763	I-DTKKKKPTSPVKPMPQNTETRTTRVRKNTGSKVAVNTEET	-----FSDNKSJKOS	1813
Qy	324	FYEFGHLYNDRSNKRGHRSKEYIVDDEDDVSRODEYENQORREDEEYQARYRSPNLARVP	383	
Db	1814	LKNNRKDLN	-----DKLPDND-----RYGGETTDSPIHHYAP	1846
Qy	384	VKPOPEOMRIHAEVSRARHRRHSVDVSLANAELDRSILRLMD	---RPSRORSVER	440
Db	1847	IEGTLY	-----CFSRMD-----SLSSLDFFDDDDVLSREKAELRKGESKDSKA	1890
Qy	441	RAAMENORSYSMEKTRERAGOSSTYQORTSNHSPPTPRSRPIL	-----DRPMRADSL	494
Db	1891	KYTCHEPSSSOQASARKAASATKHP	---VNRGSKPLQEDQPTFPOSSKDVPRGAATDE	1947
Qy	495	RKOHLLDPSAARVTKREKMETMLRNDLSL	-----DOSESVRPPRPAPKSK	543
Db	1948	KIONPAIENTPV	-----CFSRNSSLSSLDVDQENNNNETGPVRAAE-ANAOQ	1997
Qy	544	GGRMROVSLSSSEELASTPEYTCDDVLESESVSEKGDQSGKRTKSQGYVLSDSNTR	603	
Db	1998	PKRPOASGAPKSPFHEDDVPVCFSRNS	---SLSSLSID-----SSDDLRLRECIS	2044
Qy	604	SRKOKRMATYGGHSLSEDIEMWBPQ	---IKDSGVDTCCSSTTLNEHSHS-DKHPVTM	656
Db	2045	AMPKRR	---PSRLKEGEMQSPKRVGSLVLAEDLLDKIDORPSEHSLSPDSEMFDM	2100
Qy	657	OPSKDDRLIGRILLNKLRLKDSVPR	---DSGAMLGLKVVGKMTESGRILCAFTKVKR	712
Db	2101	KALQEGANSIVSSLHOAAAAACLSIQASDSDSIILSK	-----SG-----VSL	2144
Qy	713	GSILADYVGLRPDEYLENNG	---BLLOGATREEVYNNILLESKPEQVELVYSRPIGDI	768
Db	2145	GS	---PFIHLPQOEKPTFSHKGRILKGE---KSTLEK	2179
Qy	769	PRIPSTHAQLESSSSSFESQK	-----MDRPSIYTSFMSPGMLRDVPOFLSGOLS	819
Db	2180	-----KISEBNGKIGKGVKUYKSLTGKTRNSNEISSOMKQLOTUMPSISRGRTM	2230	
Qy	820	IKIMFQVGHOLIYTLGAKDLPSREDGR	---RNPYKVIYFLPDRSDKNNKRTK-TYVKTL	877
Db	2231	I	---HIPVNRSSSTSVSKGPKPLKTPASSPSEGVAATSPROTKAIVASEL	2282
Qy	878	EPKMNQTFIYSPVHREFEFRMLEITLMDQARVEESEFLELLELLELALDDEPHMY	937	
Db	2283	SPITRQTSHTSGNKGSRSGSDST	---PSRPTQ	2315
Qy	938	KLOTHVSVLLPRPSBYLPRQD	---HGESPTRLQRSKRISDSEVSDYDCEDGYGV	993
Db	2316	-----PLSRPMQSPGRNISIPGRNIGSTPNKLSQLPRTSSPSTASTK	SSGSGKM	2364
Qy	994	SDVRHNGRDLQSSITSLVPEQVMSNHGSPGSPHRYDVIGTRTWSAPAPORNEVGH	1053	
Db	2365	S-YTSPERQLSQONLS-KOTGLSKN	---ASISIP-----KESASKGLGNQNNNSNG	2409
Qy	1054	RGRATGRCHYNTISMDRHRYWMDHYSSDRDRDCEADAROPYHRSSTEQRLLERTTTRS	1113	
Db	2410	-----SNKKKVELSM-----SSTKSGSGE	SDR-----SERPALVROSTFI	2444
Qy	1114	RSSERPDTNLKRSMPSLMTGRSAP	-----PALSHRPTGSGVQVSPSSTP	1166
Db	2445	K-EAASPTLRKRKEESASESLSPSSSRPDSPTRSQAOTVPLASLDPM	SLSTHPSVOA	2501

```

QY 1161 GTGRGRQOLPOLPP--KGLTERS-----AMDIERNROMKLINKYOAGSDPLE 1208
Db 2502 GGMWR-----LPPNLSPILEYDGRPSKRKHIDARSHSESPRLPVNR---AGTWKR-- 2549
QY 1209 ODYHSKRYSGMDPHRGADVTSTKSSPDSDVSAVARSITESASRESFESTSYMSVOSERPRGN 1268
Db 2550 --EHSKH-----SSSLPRVSTWTRTQSSSITLSSASESESKAKSDE 2589
QY 1269 KRIEVTFSKMONROMGVSGNNLTKSTISGDMCSLEKNNGSSODPAVAGALGISGKKRRSS 1328
Db 2590 KHAVSVPGPRMKKNQVPTK-----GTWKRIKESIISPLNT--VSOTITSSG--AA 2635
QY 1329 IGAMNVAIVGLSRKRSRGSASOLQOTEGGAKKLRSTVORSTETLAVE--MRMMMTROASR 1385
Db 2636 SGAESKTLI-----YOMAVASRTEDVWVRLEDCTINPRGSRGPT 2676
QY 1386 ESTDGSMNYSYSECNLIIFPGVRLASDSOFSDFLDGLGPOLVGRQTLATPAMGDIQVGM 1445
Db 2677 GNTPEVIDISIEKEN---PSIKDSKDPQOSVSGSSPVOVGLENRL----- 2721
QY 1446 DKQGLEVEIIRAGLVAVKGSKTLPAPEYKVVLLDNGVCIAKKTKKTVARKLTLEPLYOL 1505
Db 2722 --NSFTLOEAPDQGTETKAGQS--PAPVAET-----GETCMAER-----TFSSSS 2765
QY 1506 LSFESPOGRVLIQTIWGDYGRMDKSPFMAVOAILDELELSNMVTIGFKLEPPPSLNDP 1566
Db 2766 SSKHSSPGFVAARVETPFNNYMPSPRKSADSTSA-----RPSQIDTP 2807
QY 1566 TSAPLITRAASOS--SLESTGTSYR 1589
Db 2808 VGSSTKKRDSKTDSTRESSGAOSPRK 2832

```

```

RESULT      8
APC_MOUSE
ID APC_MOUSE STANDARD; PRT; 2845 AA.
AC Q61315; Q62044;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ADENOMATOUS POLYPOSIS COLI PROTEIN (APC PROTEIN) (MAPC).
GN APC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 1 AND 2), AND VARIANTS.
RC STRAIN=C57BL/6J, AND CAST/EI; TISSUE=Brain;
RC MEDLINE=92263101; PubMed=1350108;
RA Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R.,
RA Luongo C., Gould K.A., Dove W.F.;
RT "Multiple intestinal neoplasia caused by a mutation in the murine
RT homolog of the APC gene."
RL Science 256:668-670(1992).
RN [2]
RN ERRATUM.
RA Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R.,
RA Luongo C., Gould K.A., Dove W.F.;
RA Science 256:1114-1114(1992).
RN [3]
RN SEQUENCE OF 1-45 FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RA Dicker F., Lambertz S., Reitmair A., Ballhausen W.G.;
RT "The murine APC gene: alternative splicing of 5' untranslated
RT region segments."
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [4]
RN ALTERNATIVE SPLICING.
RX MEDLINE=94061824; PubMed=8242607;
RA Oshima M., Suiyama H., Kitagawa K., Taketo M.;
RA "APC gene messenger RNA: novel isoforms that lack exon 7."
RT

```

RL Cancer Res. 53:5589-5591(1993).
CC -1- FUNCTION: TUMOR SUPPRESSOR. ALLOWS THE RAPID TURNOVER OF BETA-
CC CATEININ. APC ACTIVITY CORRELATED WITH ITS PHOSPHORYLATION STATE
CC ALLOWS THE DOWNREGULATION OF CYTOPLASMIC BETA-CATEININ (BY
CC SIMILARITY).
CC -1- SUBUNIT: FORMS HOMODIGOMERS AND ASSOCIATES WITH CATEININS (BY
CC SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1 (SHOWN HERE), 2, 3 AND 4; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, SPLEEN, KIDNEY, HEART,
CC LUNG, BRAIN, STOMACH, INTESTINE, TESTIS AND OVARY.
CC -1- PTM: PHOSPHORYLATED BY GSK-3B (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 7 ARM REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M88127; AAB59632.1; -;
DR EMBL; U02937; AAA03443.1; -;
DR HSSP; Q02248; 2BCT.
DR MGD; MGI:88039; APC.
DR InterPro; IPR000225; -
DR Pfam; PF00514; Armadillo_seg; 4.
DR PROSITE; PS00176; ARM_REPEAT; 1.
KW Anti-oncogene; Phosphorylation; Alternative splicing; Repeat;
KW Coiled coil.
FT DOMAIN 1 61 COILED COIL (POTENTIAL).
FT DOMAIN 125 245 COILED COIL (POTENTIAL).
FT REPEAT 1 728 LEU-RICH.
FT REPEAT 451 493 ARM 1.
FT REPEAT 503 545 ARM 2.
FT REPEAT 546 589 ARM 3.
FT REPEAT 590 636 ARM 4.
FT REPEAT 637 681 ARM 5.
FT REPEAT 682 723 ARM 6.
FT REPEAT 724 765 ARM 7.
FT DOMAIN 739 2834 SER-RICH.
FT DOMAIN 1130 1156 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 1556 1575 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 1864 1891 HIGHLY CHARGED.
FT DOMAIN 243 276 MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT VARIANT 310 410 MISSING (IN ISOFORM 3 AND ISOFORM 4).
FT VARIANT 120 120 V->A (IN STRAIN CAST/ET).
FT VARIANT 493 493 V->I (IN STRAIN CAST/ET).
FT VARIANT 797 797 Y->F (IN STRAIN CAST/ET).
FT VARIANT 1330 1330 A->T (IN STRAIN CAST/ET).
FT VARIANT 1618 1618 A->S (IN STRAIN CAST/ET).
FT VARIANT 2294 2294 G->A (IN STRAIN CAST/ET).
FT VARIANT 2496 2496 H->Q (IN STRAIN CAST/ET).
FT VARIANT 2523 2523 T->A (IN STRAIN CAST/ET).
FT VARIANT 2813 2813 T->S (IN STRAIN CAST/ET).
SQ SEQUENCE 2845 AA; 311086 MW; 145CA73CF570AA99 CRC64;

Query Match 3.1%; Score 254.5; DB 1; Length 2845;
Best Local Similarity 18.8%; Pred. No. 0.00017;
Matches 362; Conservative 241; Mismatches 692; Indels 633; Gaps 86;

QY 45 DROKKEKEKESV--LKIKER--HKAQPTQW-----PPFGSTTELNVNV 84
DB 1147 EEEHEHEEHEEPTNYSIKYNEKHVHVDPIIDSLKYADISSQKPSFSKNSAOST- 1205
QY 85 LQPOOKOP---NEKEPOTKLHOQFEMKYKMGEESSQOQOEKADPTC--GICHT 138
DB 1206 -KPEHLSSSENTANPPPNAKRONLRPSSAQRNG-----QTKG--TTCKVPSINQET 1256
QY 139 KPADGCHNCYSY--QTKFC--ARCGRVSLRSNKVMVVCNLCKQOELLTKSGAWFYNSG 195

DB 1257 -----IQTCVEDPTPIFCSSSLSSADDEIGCDQTTQTEAD-----S 1296
QY 196 SNTLOOPDOGVPPGLRNEEAPOEKKAKLHQPPQFGAPGDLVPAVEKG-----RAH 247
DB 1297 ANTLQTAEVKENDVTKSAEDATE-----VAVSONAKAKSRLOAS 1338
QY 248 GLTRQDTIKNGSGVKHIOIASDMPDRKRSVSRDQNRKRYBOSEEREDYSQVPSDGTMP 307
DB 1339 GLSSEST-----RHKKAVFSSGAK--SPSKSGAQT-----P 1368
QY 308 RSPSDYADRRSQREPOFYEERGHILNYDSNRGRHRSKEYIVDDEVDSEYERQOREE 367
DB 1369 KSP-----PEHYVQETPLVFSRCT-----SVSLDSFESRSIAS 1402
QY 368 EXGARVRS-----DENLARYPK---PQYEQMRILHAYSRAR---HERR 407
DB 1403 SVQSEPCSGAVSGIISPSDLPDSFGQTMPSPRSKTPPPPTQVAKKRVPSKPAAKR 1462
QY 408 HS--DVSLAAAELE-----DSRISILRMDRSRORSVER 440
DB 1463 ESGPKQTVANNAVQVRQVLDVDTLHFATFESTPDGFCSSSLSLASLDEFFIQKVELR 1522
QY 441 RAAMENQSYSMERTBAQCOSSYPTQRTSNHSPTPRRSPIPLDRPDH----- 488
DB 1523 IMPPVQENDNGNETESQPESENQKEVEKPDSEKLLDDSDDDDEILEECITISAMP 1582
QY 489 -----RRADSLRKQHHLDPSASVVRKTKREKMTML--RNDSSSDQSESRP--PPPR 537
DB 1583 TKSSRKAKKLAQYASKLPPRYAKRPSQILPYKLLPAQNRLOAQKHSVTPDQDVRYVCV 1642
QY 538 ---PHKSKGKGMROVSLSSSEELAS-----TPYEYSCDVELESESVSEKGSQ 585
DB 1643 EGRPINSTATSLSDLTIESPNNELATGDGVRAGIQGSEFFEKRTDIPLEGSTD--DAQ 1699
QY 586 KGR-----RKTSEGVLSL-----SNTRS--- 604
DB 1700 RGIISIVTDPDLDNKAEEGILAEICINSAMPKKGSHKPFVKKIMQVOOQASSSSGAN 1759
QY 605 -----ERQKKRMYYGHSLEEDLEWSEPOIKSGVDTCSTILNEHSID--KHPTW 656
DB 1760 KNOVDYTKKKKPTSPVKPMPQTEYRTYVRK---NTDSKVNWTEETPSDNKSKKPSLQ 1815
QY 657 QPSK-----DGDRLIGRLLNK----- 673
DB 1816 TNKAKFNEKLPNNDRKRGFTALDSPHHYPTIEGTPYCFSSNDLSLDFDDDDVLSRE 1875
QY 674 --RLKDGSVPPDSGAMLGK--VVGKMTESGRLCATTKYKKGSLADTVGHLRPGDEVLE 730
DB 1876 KAEILRKESKESDEBAKYTCRPEPNSOQAAASKSOASIKHPANRAQSKPVLOKQFPQSS 1935
QY 731 WNGRLLOGATFEFVYNNILLESKPERQVELVYSR-----PIDDIPRIPOSTHAQLESSSS 785
DB 1936 KDGDRCGAATDEKQNLATIENTP-----VCFSRNSSLSDIDID-----QENNNKE 1982
QY 786 FESQKMDRPSISYSPSPGM-----LDVPOFLSGOLSIKLMF-----DKYGHOLI 832
DB 1983 SEPTKEAEPANSGEPKPKQASGYAPKSFHVEDIPVCSRRNSSLSSLSDSDDLQGCICI 2042
QY 833 VTIIGAKDLPSREDGRPNRYVYKIFLPSRSDKNKR-----TKYVKLTLEPKWN 882
DB 2043 SSAMPKRRKRSR-----LKSESEKQSPRKVGILAEIDLTLDLQDLQRPSE 2088
QY 883 QTF-----IYSPVHRERERERLETLTDQARVREESFELGELLIE 924
DB 2089 HAVSPGSENDWKAIQEGANSTVSLHQAAAAAACL-----ROASSD--SDSLIS 2137
QY 925 LETALLDEPHWYKLTQHDVSSLPL-----PRSPYLPRLQGLSEFTRRLQSRK 974
DB 2138 LKSGISLIGSP--FHL--TPDQEKEPFTSNKGRILKPGKSLLEAKKI--ESNNKIKGKG 2192
QY 975 RI-----SDSEVSDYDCEDGVGVSYDRIHNGRLQSSLTSLVPPQVMSNHICSP-- 1022
DB 2193 KYVKSLITGKIRSNSEIS--OMKOPLPITNMPISRGRTM-----IHIPGLRNSSSSTSPVS 2247


```

Db 925 KFEDEGAFEESEETGDYEEKAETEEAEPEDEDEHVCASAKSHSPTE--DEESAKAEAD 983
QY 324 FYEPBGHNTYDSNRGRHGRHKEYIVDEDEDVBSRDEYEQ-----RREEYQARYSDSN 378
Db 984 AYIREKRESVAGSDRAEDMEDALEKEGADEQSEEADEEDKAEDAREEYE----- 1035
QY 379 LARYPVKPOPEQOMRIHAEYSRARH-----ERRHSDVSLANAELEDRISLRLMDRPSRO 434
Db 1036 -----PEKHAEDYVAAVYDKAAEAGAEQYGLTPTPQO-----LGAQSPGHE 1080
QY 435 RSVERRAAM-----ENQSYSNEMTREAOQO-----SSYQPTSNHSPPTPRRPIPLDR 485
Db 1081 PASSIHDTLPGCSSESEATASDEENREDQPEEFTATSGYOSTIEIS-----SEPTPMDE 1135
QY 486 PDMRRADSLRKOHILDPSSAAVARKTKREKMETLRLNDSSLSDQSESVRPPRPHKSKGG 545
Db 1136 -----MSTPRVAMSDETNNEETESPQOE----- 1159
QY 546 KMRQVSLSSSEELASTPEYTSODVY-ELESESVSEKSGDKGKRTSBOGLSDNTRS 604
Db 1160 -----VNTIKYESSLYS-QEYKPADVTPPLNGSESGKTDATGCKOYNASASTISPPS--- 1211
QY 605 ERQKKRMYTGGHLSLEEDLEWSEPOIKDSGVDTCS-----STLINEHSHSKDHPVTWQPSK 660
Db 1212 -----SMEDD-KFSKSLALRDA---YCSFKASTTLDIKDSISAVSEKVS 1255
QY 661 DGDRLIGILLRLKDKSVPRDSCAMGLKAYVGKMTESGALCAFITVKKGLADYTG 720
Db 1256 -----SPSLSPSPSPLEKTPLEGRSVNLSLPN----- 1284
QY 721 HLRPGDEVLEWNGRLLOQATFEENVYIILSKPEPOVELVSRPGIDIPRIDPSTHAOLE 780
Db 1285 -----ELKVAEAEVAVSEYVQOEYVE 1307
QY 781 SSSSSFEQSKMD--RPSISVTSMPSPMLRDVPOFLSGOLSIKMLFDKYHOLIVTILGA 838
Db 1308 EHCASPEDKTLEVVSPSQSVT-----GSAGHTPYQOVSPT--DEKSSH----- 1347
QY 839 KDLSREGRGRNRPVYKTYF-LPDSDKNKKRT-----KTVKKTLEPKNNQTF 885
Db 1348 -LPLEVLEKP--PAVPVSPFEESDAKDENERASVSPMDPEVPDESPIKIVLSPLSPPL 1403
QY 886 IYSPVHREPREBMLLEITLMDQARVREESEF-----LGEILIELETLALDD- 932
Db 1404 IGSEBAYESF-----LSADDKASGRGASPEEKSCKGSGSDQVSPVEMTSLTYQDK 1457
QY 933 -----EPHMYKLOTHDVSSLPRLPRSPYLRROLHGESPTR-----RLQ 971
Db 1458 QEGKSTDFAPIKEDGOKKTDVDEAMS--SOPALALDERKLDVSPTOIDVQFGSKFED 1516
QY 972 RSKRISDEVDYDCED-GVGVVSD-YRANGRDLOSSLYS-----PEQVMSNHGSPS- 1023
Db 1517 TKMSTSECTVSDKSAVPDEGVAEETYSIMEGVASVTSVATSSPEP--TTDDVSPSL 1574
QY 1024 -----GSPHVRVYIGRTSRMSPSAPPPQRNVNEGHRATGYNNTISRDRIHVMDDHS 1079
Db 1575 HAEVSPRISTEV-----DDSLIS 1591
QY 1080 SDRDRCEADROPYHRSRSTEQRLBERTTTRSSRSEPRDNLNKMSP--SLMTGRSAP 1137
Db 1592 V-----SVUQTPTTEQTEMSPSKKECPRPMSTISPDSPKTAKSRT 1633
QY 1138 PSPALSRHPTGSV---QTSPTSPTGGRGRQLRQLRPKGT-----LERSA 1182
Db 1634 PVQD-HREQSGSMSTIEQOESPEQSLAND-FSRQSPDHPTVAGVLHTENGTEVEDYDPS 1691
QY 1183 MDIEERNQMLNKQYAGSDPRLQEDYHRSRGWMDHGRADVSTKSDVDYDVSA 1242
Db 1692 SDMODSSLSHKI-----PRMEBSYTO-----DNDLSLLIS 1722
QY 1243 VSRITSSASRFSSTYSVQSERPRGN-----RKISVFTSKMKNOMQVSGKNILTRS 1293
Db 1723 VSOV-EASPTSTSAHTPSQIASPLQEDTLSDVAPPRDMSLVASTLSEKQVSLGCKLSP 1781

```

```

QY 1294 TTSIGDMCSLEKNDGSDPAVAGALGTSGKKRS-----SIGAKNVAIVGLSRKRSASOL 1349
Db 1782 SDIS-----PLTPRESSPLYSTESDSTSAVKEKTATCHSSSSPPDAAASAPYGERASVL 1837
QY 1350 SQT-----ECGCK---KLRSYVORSTETGLAVEMRMNTROASREST 1388
Db 1838 FDMQHHLALNRDLSTPGLKEDSGCKTPGDSYAVQKEEL-----TRSPDEEDY 1887
QY 1389 DGSNNT-----SSEGNLIFPGV-----RLASDSQSFDLGLGPAQLVGRQTATPAMG 1438
Db 1888 D--YESYEKTTKRTSDVCGYVYKERTTKSPSDSGYS-----YETIGK-TTKTPEDG 1936
QY 1439 DIQVGMDDKKGOLEVEILIRARGLVKPGSKTLPPYVYVLLDNGVCIAKKTKYARKTL 1498
Db 1937 DYSEIIEKTRTPEE---GGYSYDISEKTTSPPEVSGY-----SYETKERSRL 1984
QY 1499 EPLYOQLSFEESPOG 1514
Db 1985 DDLSN---GYDDSEDG 1997

RESULT 10
NCRL_HUMAN
ID NCRL_HUMAN STANDARD: PRT; 2440 AA.
AC 075376; Q9UPV5; Q9UQ18;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NUCLEAR RECEPTOR CO-REPRESSOR 1 (N-COR1) (N-COR).
GN NCOR1 OR KIA1047.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Petal brain;
RX MEDLINE=98393736; PubMed=9724795;
RA Wang J., Hoshino T., Redner R.L., Kajiyaya S., Liu J.M.;
RT "ETO, fusion partner in t(8;21) acute myeloid leukemia, represses
RT transcription by interaction with the human N-COR/msin3/HDAC1
RT complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10860-10865(1998).
RN [2]
RP SEQUENCE OF 782-2440 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9937452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K.-I., Hirosewa M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
RN [3]
RP SEQUENCE OF 974-2440 FROM N.A.
RX MEDLINE=99375328; PubMed=1044336;
RA Nagaya T., Chen K.-S., Fujieda M.,
RA Horwitz K.B., Lupski J.R., Seo H.;
RT "Localization of the human nuclear receptor co-repressor (hn-COR) gene
RT between the CM1A and the SMS critical regions of chromosome
RT 17p11.2.";
RL Genomics 59:339-341(1999).
CC -|- FUNCTION: MEDIATES THE TRANSCRIPTIONAL REPRESSION ACTIVITY OF SOME
CC NUCLEAR RECEPTORS BY PROMOTING CHROMATIN CONDENSATION, THUS
CC PREVENTING ACCESS OF THE BASAL TRANSCRIPTION.
CC -|- SUBUNIT: FORMS A LARGE COREPRESSOR COMPLEX THAT CONTAINS SIN3A/B
CC AND HISTONE DEACETYLASES HDAC1 AND HDAC2. THIS COMPLEX ASSOCIATES
CC WITH THE THYROID (TR) AND THE RETINOIC ACID RECEPTORS (RAR) IN THE
CC ABSENCE OF LIGAND.
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -|- DOMAIN: THE N-TERMINAL REGION CONTAINS REPRESSION FUNCTIONS THAT
CC ARE DIVIDED INTO THREE INDEPENDANT REPRESSION DOMAINS (RD1, RD2

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF044209; AAC33550.1; -
CC EMBL; AB028970; BAA82999.1; -
CC EMBL; AB019524; BAA75814.1; -
CC MIM; 600849; -
CC InterPro; IPR001005; -
CC Pfam; PF00249; myb_DNA-binding; 2.
CC PROSITE; PS50090; MYB_3; 1.
CC Nucleic protein; Transcription regulation; DNA-binding; Repressor;
CC

```

Query Match	3.18;	Score 253;	DB. 1;	Length 2440;
Best Local Similarity	18.48;	Pred. No. 0.00016;		
Matches 329;	Conservative 618;	Indels 614;	Gaps 76	

Db	336	SKTRREYKOP	-----PEIKRQROQERFQVYRGAGLSATIRASEHSEITIDG	386
Qy	267	-SDMPSDKRSPVSRDQNRREYEGOSEEREDYSQVYPSGTPMRSPSDYADR	-----SQR	320
Db	387	LSQENNEKQKROLSTVPRMFDAEQR	-----VKFLNNGLMEDPKYVKDQFMNWTDH	443
Qy	321	EPQYEEFGLNLTNSNRGHRHSKEYLVDEDDVSRD	-----EYEQRRESEYQARYR	374
Db	444	EKEJFKD	-----KFTONRKNFLLASLYERKSVPCVLYLYLLTKNENKALVR	492
Qy	375	-----SDENLARYPVKPOPYEOMKITHAEVSHAREHRHSDVSLANALEDSTRISLL	426	
Db	493	RNYGRRGRNQOIAN	-PQOEKVEKEEDKAEKTEKKEEKDEEE--KDEKEDSKENTWK	549
Qy	427	RMDPRSORVSERR	-----AMENORSYMER--TPEAGOSSYPORTNSHSP	474
Db	550	EKDKIDGTAEEETEEREQATPRGKRTANSQGRKRGITRMTMEAAASAASAAAEDEPRP	609	
Qy	475	TPRSPIDLRDMRAASLRKO	-----NHLDPSAVKTRKREKMETILRN	520
Db	610	PLPPEPELTSEPVTETSKRTEEMEYAKGLVHEGRNMAALAKMGTKEACCKNFYNY	669	
Qy	521	-----DLSLSDQSESVRPPRPKSKKGKMYOVSLSSSEELASTPEYISCDVLE	574	
Db	670	KRHMNDNL	-----QKHOKTSRKPRERDVSQCEASVATSAOEDIEDIAS	717
Qy	575	SESSEKGDQKGRKTS	--QGVLSDSNTSEKOKKMYGGHLEEDL	622
Db	718	NEEENPE	-DSEVEAYKPSDESPENATSRNTPEAVLEPTTETAPTSPSLAVPTKPAE	776
Qy	623	-EMSEPOIKDS	-GVDCSSTLTN-----EESHSHKHNVTMOPSKD	661
Db	777	DSVETQVWDVSTISAEIADQMDVDQOHSABEGVCDPRATADSVDEVYRPNENASKY	836	
Qy	662	-GDRLIGRLILNKLKLDGVPFRDGSAMLKLVYGM	-----TESGRCAFTIKYKGSJ	715
Db	837	EGGNTKERDL	-DRASEXVEPRDEDLVVAQOJIAQPREROSDMSDATOS	884
Qy	716	ADTVGHLRQDEVLKNGRILGQATGEYVYNILLESKP	-----EPQVEIYVSRPIG-----DI	768
Db	885	-----ADEVDDEP	-----ERQMRPMDSKPSILNPTGSLIVSPSLKPNPDL	927
Qy	769	PR	-----IPDST-----NAQLESSSSPESQKMDRPT	796
Db	928	POLOHAAVTPRWVSCSTPCNIPIGTYSVGYALVORNIKAMNESALLBEOHOBODILEC	987	
Qy	797	-SVTSP	--MSPGMLRDVPOJISQDLSIKLMDKYGHOLIVILAKKDLSPNEDGPRN	851
Db	988	RSSTSPGTSKSPNREWEVLO	-----PAVNOILITMLPEGVRLPTTPRTPRP	1034
Qy	852	PVYKIFELPRSDKNKRRKTYKTLERKWNQTFIYSPVHRERFERMELITLMDQAVR	911	
Db	1035	PLIP	-----SSKTT-----VA	1045
Qy	912	EESEFPLGELLIELETTALDDERHWYKLOTHDVSSPLPRSPY	-----LPRQOLHGE	964
Db	1046	SEKSPFL	-----MGGISQIGTPQTY-LTSHNQSXYQETPRKPSVGSISLGLPRQOESAK	1098
Qy	965	SPT	-----RLQORSKRISDSEYSDCEDGCVGVSDDYHNG	1005
Db	1099	SATLPYIKOEFEFSPRSONOPE	-----GLVIRAOHGEVVGKSTAGALQDGSITTKGPT	1150
Qy	1006	STLSEVPQVWSSNHCSPGSPHNVAVLIGRTSMGSPSAPRPORANEOGHRGTRATGHNTI	1065	
Db	1151	SKISV	-ESISPLKRSITOGTRA-----LPQIGIPTEALVKG	1187
Qy	1066	SRMDRHRVMDHYSSDRDQCEADR	-----QVHRKSGSTQKRLLETT	1110
Db	1188	SRMP	-----IEDS-SPEKGRE-EAASKGHVITYEGKSGHILSYDNIKMAREGTRSPTAHEI	1241
Qy	1111	TRSSSERPPTNLMRSMPLMTGSGAPSPALSHSP	-----TGSV-QTSP	1156

Db 1242 SLKSYSEBECNIKQMSMRSPVAPLEGICALRPRGSPHSDKERTVLSGIMQGP 1301
 QY 1157 SSTGTGRGRLQPLPKGLTLERSAMDIEENRQMLNKRYQVAGSDRLRQDYHSKR 1216
 Db 1302 RATESFEEDGLKP-----KQIRRESPIT-RAFGAIT 1333
 QY 1217 SCMPHRCADVSTKSSSDVDVSAVSTSSASRFSSTYSVMSEPRGRNKSIFTS 1276
 Db 1334 KG-KPYDITTIKEMG-----RSIHEIPRQDILLQESKRTPEVOSTRIIEGSIQGP 1387
 QY 1277 -KMONRGVSGKN-----LTKSTISGDMCSLE-----KNDGSDPTAVGAL 1318
 Db 1388 IKFNN-----SGOSAIKHNVSLITPGSKLSKGMPLLEVPENIKVERGKYEDVAGE- 1442
 QY 1319 GTSCKRRSSIGAKNVAIVGSRKRSASOLSOTEGGKKLSTVQSRFETGLA----- 1372
 Db 1443 -----TVRSRHTSVVS---SGPSVLSTLHEAKQAQISPGIYD 1478
 QY 1373 -----VEMRNMTRQA-----SNESTDGSMNSYSSEGNLIFPGV 1406
 Db 1479 TSARPTVSYQNTMSRSGSPMNRSTDVITPRKSTNHEKSKTLTPQRESIPAKSPVGP 1538
 QY 1407 -----RLASDQSPDFDLGPAQLVGRQTLATP--- 1435
 Db 1539 DPVYSHSPDPHHRGSTAGEVYWSHLPTQLDPAMPFHRALDPAAMAAYLFQRLSPTPGP 1598
 QY 1436 -----AMGDLYGVAMD-----KKGLEVEITRARGLYVKGSKTLPARY 1474
 Db 1599 SQYLYAMENRQTLINDYITSQMQVNLRPVARGL--SPREDPLGLPY 1646

RESULT 11
 NCR_MOUSE STANDARD: PRT: 1453 AA.

AC P30415; 01-Apr-1993 (Rel. 25, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE NK-TUMOR RECOGNITION PROTEIN (NATURAL-KILLER CELLS CYCLOPHILIN-RELATED PROTEIN) (NK-TR PROTEIN).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93133824; PubMed=8421688;
 RA Anderson S.R., Gallinger S., Roder J., Frey J., Young H.A.,
 RA Orcalido J.R.;
 RT "A cyclophilin-related protein involved in the function of natural killer cells";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
 CC -1- FUNCTION: COMPONENT OF A PUTATIVE TUMOR-RECOGNITION COMPLEX.
 CC -1- INVOLVED IN THE FUNCTION OF NK CELLS.
 CC -1- SIMILARITY: CONTAINS A CYCLOPHILIN-LIKE PPIASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L04289; AAA37500.1; ALT_TERM.
 DR HSSP: Q27450; 1A58.
 DR MGD: MGI:97346; Nkr.
 DR InterPro: IPR002130;
 DR Pfam: PF00160; pro.isomerase.1.
 DR PROSITE: PS00170; GSA_PPIASE_1; 1.
 DR PROSITE: PS50072; GSA_PPIASE_2; 1.
 DR Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
 FT DOMAIN 1 176 PPIASE, CYCLOPHILIN-TYPE.

FT DOMAIN 222 241 ARG/LYS-RICH (BASIC).
 FT DOMAIN 422 459 ARG/LYS-RICH (BASIC).
 FT DOMAIN 964 1003 ARG/LYS-RICH (BASIC).
 FT DOMAIN 198 273 ARG/SER-RICH.
 FT DOMAIN 468 565 ARG/SER-RICH.
 FT DOMAIN 658 812 ARG/SER-RICH.
 FT DOMAIN 1303 1453 ARG-SER TANDEM REPEAT-RICH.
 SO SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;
 Query Match 3.0%; Score 248.5; DB 1; Length 1453;
 Best Local Similarity 18.0%; Pred. No. 0.00014;
 Matches 295; Conservative 185; Mismatches 570; Indels 591; Gaps 62;
 QY 14 TPASQPPPEPMDLSHLEERKILLAVDRCK-----EE 51
 Db 158 TDAASRPYADVRYIDCVLA---TKLTQVFEKKRRKPTCEGSDSSRSSSSSESSES 214
 QY 52 EKESQVYKIKEENKAOPTQWFFSGITELVNNVLQPOQKQNEKEPQTKLQOEFMYKEQ 111
 Db 215 EFERETIR-RRHKRRRK-----YHAKKRREKSSSEPRRKRTVSPGYSER 262
 QY 112 VKMGEESQOQOEGKGPAPICGICHTKTFADGCGHNCYCOTKFCARCGVSLRSKNV 171
 Db 263 -SVNNEKRSVDSNTRKRPV----- 288
 QY 172 WVCN---LCRQKQELTKSGAMFYNSGNTLQPDQKVPRLRNEAPRQEKAKLHQEPQ 228
 Db 289 PVENRFLRLRDMFAIT-----VEPDNIP-----DVAVVSDQK----- 323
 QY 229 FQGAPGDLVPAVER-GRAGQLTRQDTIKNGSGVYKHQIASDMPDRKRSVSRDQNR 287
 Db 324 -----PSVSKSR-----KIKGKGTIRH-----TPRSHSHESKDDSDSE 360
 QY 288 ---EQSEREDYSQYVPSDGTMPSPSDYADRSCOREQ-----TYEEFGH 330
 Db 361 TPWHKKEEMQRLRAYRPPSGCKWMSKGDLSDPCCSRWDEKSLSORSRWSYNGYSDLS 420
 QY 331 LNRDNRGRGRHSHKEYIVDEDEVESNDEYEROREEYQARYSD-----PVL--AR 381
 Db 421 ARHSDGHHKHKRKKF-----KHKAKKQKHCRRHROTCKRRLVMPRLPSR 469
 QY 382 YP--VKPQYEEMQRIHAESYRARHERH---SDVSIANAELDESISILRMDRPSQ 434
 Db 470 SPTHRMKSQCRER-RBRASSSSSHSKRDMWSKSDQDDGASVSHSDYRSKSHSD 528
 QY 435 RSVSERRAMENQSYSMERTREAGQSSYTPQTS----- 469
 Db 529 SRGSSRRRAVSKSSSRSLNRSKSSSSSRSGPRRTSISPKKPAQLSENKPVTEPLRPSV 588
 QY 470 -----NHSPTPRRSPILPD-RPDMRRADSLR-KONHLDPSSA-- 505
 Db 589 QNGNVLQVPAENIPIVPLSDSPPRSKMGQKPKPSTYERIDEMKAKTTHLLPVQSTY 648
 QY 506 ---VRKT-----KREKME-----TMLR 519
 Db 649 SLTIKATLVSSSYHKREKPEESDGSAYSKYSDRSSGSSGSKSSRSRSTYTSR 708
 QY 520 NDSLSQSSQSVPRPPRPHKSKGKMKQVSLSS-----EELAS 561
 Db 709 SRSLPTSRSLR-RSPSSSRSHSPNKYSQSHSRSSSYSVSSDDGRRAMFRNRKKS 767
 QY 562 TPETSSDDVLESESVSEKSDQKGRKKTSEQGLVLSNTRSEKQKRMYYG 614
 Db 768 HKRHSRSEKTLHSHKYVGRKSSR-HRKYSKSSSLDYTSDSDSHVQVTSAPKEKQ 826
 QY 615 -----GHSLIEDY---EMSEPOLKSDGVDTCCSTTLNEHSHDKHYVTQPSKDG 662
 Db 827 KVELNDKQGGREGKPRKPMWECPRSKEN-----SEDS 862
 QY 663 DRLIGRILLNKLKDGSVPRDSGAMLGKVVGGKMTESGRLCAFTYKKSGLADTVGHL 722
 Db 863 -----RDDSVSKGKNCAGSKWDSSENSSEODVTKSRKSD-----P 896

```

OY 723 RPDGEVLEWNGRLLOGATFEFYNNIIESKPPROYELVYSRPIGD---IPRIIPSTHAOL 779
DB 897 RRGSEKEGEA---SSDSSEVGOSHIRAKKPAKPTSTFLLPESDCAWMSRRQSSASES 953
OY 780 ESSSSFEQKMDRPSISVTPSPMGLRDVPOFLSGOLSIKLMFDKVGHLLVTLIGAK 839
DB 954 ESSCENL-----GNINGEPQ-----K 969
OY 840 DLPSEDEGRPNRYKITYFLPDRSDKNKRTKVTKLTLEPKNMOTFLYSVHRREFRERM 899
DB 970 QKHSKDLKGDH---TKRAREKSKAKDKKHKAPKRRQARHWOP----- 1010
OY 900 LETIMDQARVAREESEFELIETLLELALDDEPHMYKLOTHDVASLPLPRPSYLLPRR 959
DB 1011 -----PLEFG--DDE-----EE 1020
OY 960 QLHGESPTFRRLQSRKRISEVSDYDCE---DGVGVSDYRHNGRDLQSLVSYPEQ--V 1014
DB 1021 EKNKGQVTDOPKEKRHYSEK-----CEAVKDGIPNVEKTCDESSPSKPKKGTLEDDPL 1074
OY 1015 MSSNHCSPGSPRNVYIGRTSRWSPSAPPPOANNVGC-----HRG 1055
DB 1075 AEGGH-DPSSCAPLKEVEDNTASSPSA---QHLEHHPGGGEDVLOTDDMETCTPDRT 1130
OY 1056 TRATGHYNTISRMDRRVMDHYSSDRDRD-CEAADRQPYHRSTRTEORPLEETTRSR 1114
DB 1131 SPRKG--EYVSPPLANIRHLSPEVNIIPEODECMA-----HPRAGQESSMSKTLGE 1182
OY 1115 SSERPTNLMRSPSLMTGSAAPPSPA---LSHSHPRTG---SVQTSPTSGTGTRG 1166
DB 1183 SGVKQOSSSVTSPVESIGKEGAEKQMLNLTWKPLQGGNLVSSTATSSLDVYKAL 1242
OY 1167 ROLPOLPKKTLERSAMDIERN-----ROMKLN---KYKOVASDPRLEODY 1211
DB 1243 STYPEKPGOL---REIKSKNKVRRGSLFDEYKRTARLARRRNDQSSDDQTPSRDG 1298
OY 1212 HSKYRSGMDPRGADIVYTKSSSDVDVAVSTSSAPRSSSTSYMSVQSEPRGNKI 1271
DB 1299 DSGRS---PHRSRSEKSETSRHRTRS--VSYSHRSRSRSTSYRSRSR----- 1348
OY 1272 SVTTSKQNMQMGVSGKNLTKSTISGDMCSLEKNDQSOTAVAGALTGSGKKRSSIGA 1331
DB 1349 -----SDMTSRGTRSRSS-----YSPHSRHSRSSSR 1378
OY 1332 KNAIYGLSKRSASQLOTEGGGKKLRTVQSTETGLAVEMRMNMTROASRSTGCS 1391
DB 1379 SRSSSYLHRSRRTYDYYSRSRSRSGRSDSYHRG-----RSYNRRSRSGR 1428
OY 1392 MNTSYSEGNLIFPGVRLASDS 1412
DB 1429 SYGSDSESDRSYSHRSPSES 1449

```

```

RA Otaido J.R.;
RT "A cyclophilin-related protein involved in the function of natural
RT killer cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
RN [2]
RP REVISIONS:
RA Anderson S.K.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF A PUTATIVE TUMOR-RECOGNITION COMPLEX.
CC INVOLVED IN THE FUNCTION OF NK CELLS.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE VIA ITS N-TERMINUS.
CC -1- SIMILARITY: CONTAINS A CYCLOPHILIN-LIKE PPIASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
CC EMBL; L04288; AAA35734.2; -.
CC EMBL; AF184110; AAD56402.1; -.
CC PIR; A47328; A47328.
CC HSSP; O27450; 1A58.
CC MIM; 161565; -.
CC InterPro; IPR002130; -.
CC Pfam; PF00160; PRO_Isomerase; 1.
CC PRINTS; PR00153; CSA_PPIASE; 1.
CC PROSITE; PS00170; CSA_PPIASE_1; 1.
CC PROSITE; PS0072; CSA_PPIASE_2; 1.
CC Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
CC PPIASE, CYCLOPHILIN-TYPE.
FT DOMAIN 1 176 ARG/LYS-RICH (BASIC).
FT DOMAIN 219 240 ARG/LYS-RICH (BASIC).
FT DOMAIN 421 457 ARG/LYS-RICH (BASIC).
FT DOMAIN 970 1010 ARG/LYS-RICH (BASIC).
FT DOMAIN 194 244 ARG/SER-RICH.
FT DOMAIN 466 574 ARG/SER-RICH.
FT DOMAIN 664 814 ARG/SER-RICH.
FT DOMAIN 1311 1348 ARG-SER TANDEM REPEAT-RICH.
SO SEQUENCE 1462 AA; 165676 MW; D98A1147763EF527 CRC64;

Query Match 2.9%; Score 244; DB 1; Length 1462;
Best Local Similarity 19.5%; Pred. No. 0.00023;
Matches 268; Conservative 169; Mismatches 501; Indels 434; Gaps 62;

OY 10 RPAETPAASOPPPPEMPDLSHLTBEERKILLAVMDROKKEEKEQSVLTKEENKAQPT 69
DB 299 RQMPVYTAE---PEPKIPDVAPIVSDQK-----PSVSKSGRIKRGRT 338
OY 70 QWPFSGITELVNNVLOPOQOPNEKE---PQTKLHQPEMYKEQYKMGESQOQOEQ 125
DB 339 -----IYHTPPRRSRSCSESDDDDSSETPH---WKEEMQRLRAYRPPSGEK 382
OY 126 --KGDAPTGICHKTKFPADCGCHNCSTCQTKFCARCGGRVSLRNKMYWCNLCRKQOEI 183
DB 383 WSKGD-----KLSDCCSR-----W-----DERSL 402
OY 184 LRTSGAMFYN-----SGSNLLOPDKVPRGLRNEARPOEKKAKLHEOPPOFG- 231
DB 403 SQRSSWSYNGYISDLSTARSHGHHKKRKEKKYK---HKKKGKKKHCRNKHQTKKRI 459
OY 232 -APGDLSVPAVEKGRAGLRTQDTIKNGSVK--HQIASD--MPSDRKRPVSVDQNRXY 287
DB 460 LIPSDIESSKSTTRMKSSCDRERRSSSSLSHSSKRWMSKSDKVQSSLTGSSRDSY 519
OY 288 EQSEERPDYQYVSDQTMRSRPSDYADRSQRPQFEEEGHNLVYDSNRGHNSKEY 347
DB 520 RSKHSQSYSGSSRSRTAKS--SSHSRSRSK-----SSSSSGHR--KRA 563
OY 348 IVDDEVDSEYREYRORREERYQARYSDPMLARYPV-----KROP----- 388

```

```

Db 564 SKSPKRTASQLESENKPVKTEPLRATMAQNNVNVQPVVAENIPVILPDSPPPSRMKPCQ 623
QY 389 -----VEEQRIHAHEV-----SARHERHRSDVSLNAELEDNR 422
Db 624 KPMKPSERIOEMKKTTHLLPIQSTYSLANTKETGSSSYKRE-----KNSEDSQST 677
QY 423 ISLRMDRPS-----RORSVSERRAMENORSYMERTEACQSSYPQ-----466
Db 678 YSKY-SDRSSSSSPRSRSR-----SRYSRYTTRSLASSHSRSPSSRSRSHN 730
QY 467 RTSNHSPTPRRSPPLPRDPMRA-----DSLAKOH-----ID 501
Db 731 KYSHSQSSRSSYTSISSDGRRAKRLRSSGKKNVSHKSSSEKTLHSKYVKGND 790
QY 502 PSSAVRKTREKMETMLNDSLSQSSVPRPPPHKSKGGMKROVLSSEELAS 561
Db 791 RSSCYRKSESRSLL-----DYSSDSESSVQATQSAQEKQKQ-----QMERHNNKEKNR 842
QY 562 TPPEYTSQDVELESESVSEKGDQKRRKTSQGYLSDSNTRSEKQKRMYYGSHLEED 621
Db 843 G-----EKKSKSEBECPHSKKRTIKEN--LSHLRNGSKPKRKNVAGS-----883
QY 622 LEMSPQIKDQSVDTCSSTTLNEHSHDKHPYTWQPSKQGRLLGRILL-----NKRLK 676
Db 884 -KMPSESENSEKRV-----TKNSKNDSPSSDKK--EGEATSDSESESEIHKVKTPTKSSST 937
QY 677 DGSVPRDQSGAMLGKLV-----VGKMTESGRLCAPITVYKKGSLADTV-GHLRPGDEVY 729
Db 938 NTSLEPDDGAMKSKKQRTSTSDSESCSSSENNRCKPKQHKHKGSKNLKREHTKKYKEKL 997
QY 730 E-----WNGRLQAGATFEVYNN--ILLESKPEPOVELVVRPIDIPI 769
Db 998 KKKKKKKKAPKQKAFHMQPPLERGESEEDIDKQVQSEKKEKVSE-----NNE 1049
QY 770 RLPDSTHAQLESSSSFSQKMDRPSISYTSQSPGMLDVPQFLSGQSLIKIMEKRVGH 829
Db 1050 TIKDNI--LKTEKSEEDLSGKHDTVYSS-----DLQFTKDSKLSISPTALMT 1098
QY 830 QLIYVIL-----GAKDLPREDGRPRNPVYKIFYLPDRSDKKRRTYKTKTL 877
Db 1099 EENVACLONIQHVESVPGNVEDVLOTDDNME-----ICTPDRSSPAK-----VEET- 1145
QY 878 EPKWNQTFYVHRFRERERMLETTL-WDQAR-----VREE-----ESEFLGET-----921
Db 1146 -----SPLGNARLDTPDINIVLKDMATEHQAELVAGQESSMSKVLGEVKGD 1195
QY 922 -LIELETA-----LLDEPHWYKIQTHDVSLLP-----948
Db 1196 SSSASIASAGESTGKKEVAEKSQIMLIDKK--WKPLQGVGINLAAPMAATSSAIVEKVLTT 1253
QY 949 LPRPSPYLPRQLHGESPTR-----RLOKSKRISDSEVSDYDCEDGVGVVSDX 996
Db 1254 VPEMKPQGLRIETKKNKVPKPSLFEVAKTARLNRRPNQESS-----SDE 1300
QY 997 RHNGDLOSSTLVEQVWSSNHCSPGSPHARVDVIGTRTSMSPAPPORNVEQCHRG 1056
Db 1301 QTPSRDDDSQSSS-PSRBSK--SETKSRHRTSRVSYSHSRSSSSSYSRSYSR 1356
QY 1057 RATHYNN--TISRMDRHRVMDHYSSDRDCEAADROPYHRSRTEORPLERTTTRS 1113
Db 1357 RSRGWYSKGRTRSSSSYSYKSHRTSRSRSS-SSSSYDPHSRSTYTDYSYRSR 1415
QY 1114 RSSEPDNINMSPSLMTGRSAPSPALSRSH-----PRGSSVQTSPOST 1159
Db 1416 R-SQKSD-----SYHGRKSYNRRSRSCRSYGSDESDESRYSHHRSPSES 1458

```

```

DT 01-OCT-1994 (Rel. 30, last annotation update)
DE MICROTUBULE-ASSOCIATED PROTEIN 2 (MAP2B) [CONTAINS: MAP2C].
GN MAP2.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN (1)
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=90251471; PubMed=2339070;
RA Kindler S., Schwanke B., Schulz B., Garner C.C.;
RT "Complete cDNA sequence encoding rat high and low molecular weight
RT MAP2."
RL Nucleic Acids Res. 18:2822-2822(1990).
RN (2)
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=91060576; PubMed=2174050;
RA Kindler S., Schulz B., Goedert M., Garner C.C.;
RT "Molecular structure of microtubule-associated protein 2b and 2c from
RT rat brain."
RL J. Biol. Chem. 265:19679-19684(1990).
RN (3)
RP SEQUENCE OF 1-151; 1515-1694 AND 1726-1861 FROM N.A.
RX MEDLINE=90221819; PubMed=2326166;
RA Doll T., Papadimitrakopoulou A., Matus A.;
RL "Nucleotide and amino acid sequences of embryonic rat MAP2c."
RN Nucleic Acids Res. 18:361-361(1990).
RN (4)
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=89365159; PubMed=2770869;
RA Papadimitrakopoulou A., Doll T., Tucker R.P., Garner C.C., Matus A.;
RT "Embryonic MAP2 lacks the cross-linking sileam sequences and
RT dendritic targeting signal of adult MAP2."
RL Nature 340:650-652(1989).
RN (5)
RP SEQUENCE OF 1695-1725 FROM N.A.
RX MEDLINE=94110302; PubMed=8282767;
RA Doll T., Melchiorner M., Riederer B.M., Honegger P., Matus A.;
RT "An isoform of microtubule-associated protein 2 (MAP2) containing
RT four repeats of the tubulin-binding motif."
RL J. Cell Sci. 106:633-640(1993).
CC -1- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES.
CC -1- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF MAP2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. MAP2C, THE LOW MOLECULAR
CC FORM OF MAP2, LACKS THE CENTRAL DOMAIN OF MAP2A/B.
CC -1- DEVELOPMENTAL STAGE: MAP2C IS EXPRESSED DURING EMBRYONIC BRAIN
CC DEVELOPMENT AND UNTIL POSTNATAL DAY 10. MAP2B IS EXPRESSED
CC THROUGHOUT BRAIN DEVELOPMENT.
CC -1- SIMILARITY: CONTAINS 3 OR 4 TAU/MAP REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51842; CA36135.1; -
DR EMBL: X17682; CA35667.1; -
DR EMBL: X71487; CA350588.1; -
DR PIR: S07887; S07887.
DR PIR: S10003; S10003.
DR PIR: A37981; A37981.
DR InterPro: IPR001084; -
DR Pfam: PF00418; tubulin-binding; 4.
DR PROSITE: PS00229; TAU_MAP_1; 3.
DR Microtubules; Repeat; Alternative splicing; Calmodulin-binding.
KW DOMAIN 1454 1474 CALMODULIN-BINDING (POTENTIAL).
FT

```


AC P22793; (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 40, Last annotation update)
 DE TRICHOHYALIN.
 GN THH.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93260018; PubMed=7684041;
 RA Fietz M.J., McLaughlin C.J., Campbell M.T., Rogers G.E.;
 RT "Analysis of the sheep trichohyalin gene: potential structural and
 calcium-binding roles of trichohyalin in the hair follicle."
 RT J. Cell Biol. 121:855-865(1993).
 RN [2]
 RP SEQUENCE OF 1016-1549 FROM N.A.
 RC STRAIN=MERINO-DORSET HORN X BORDER LEICESTER; TISSUE=Wool follicles;
 RX MEDLINE=90130632; PubMed=2298812;
 RX Fietz M.J., Presland R.B., Rogers G.E.;
 RT "The cDNA-decoded amino acid sequence for trichohyalin, a
 differentiation marker in the hair follicle, contains a 23 amino acid
 repeat."
 RT J. Cell Biol. 110:427-436(1990).
 RL J.
 CC -1- FUNCTION: INTERMEDIATE FILAMENT-ASSOCIATED PROTEIN THAT ASSOCIATES
 IN REGULAR ARRAYS WITH KERATIN INTERMEDIATE FILAMENTS (KIF) OF THE
 INNER ROOT SHEATH CELLS OF THE HAIR FOLLICLE AND THE GRANULAR
 LAYER OF THE EPIDERMIS. IT LATER BECOMES CROSS-LINKED TO KIF BY
 ISODIPEPTIDE BONDS. IT MAY SERVE AS SCAFFOLD PROTEIN, TOGETHER
 WITH INVOLUCRIN, IN THE ORGANIZATION OF THE CELL ENVELOPE OR EVEN
 ANCHOR THE CELL ENVELOPE TO THE KIF NETWORK. IT MAY BE INVOLVED IN
 ITS OWN CALCIUM-DEPENDENT POSTSYNTHETIC PROCESSING DURING TERMINAL
 DIFFERENTIATION.
 CC -1- SUBUNIT: HOMODIMER (PROBABLE). TWO FORMS MAY BE PRODUCED BY
 CC -1- ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE HARD KERATINIZING TISSUES SUCH AS
 THE INNER ROOT SHEATH (IRS) OF HAIR FOLLICLES AND MEDULLA, AND IN
 THE EPITHELIA OF THE TONGUE, HOOF AND RUMEN.
 CC -1- DOMAIN: CONSISTS OF NINE DOMAINS. DOMAIN 1 CONTAINS TWO EF-HAND
 CALCIUM-BINDING DOMAINS. DOMAINS 2-4, 6, AND 8 ARE ALMOST
 ENTIRELY ALPHA-HELICAL, CONFIGURED AS A SERIES OF PEPTIDE REPEATS
 OF VARYING REGULARITY, AND ARE THOUGHT TO FORM A SINGLE-STRANDED
 ALPHA-HELICAL ROD STABILIZED BY IONIC INTERACTIONS. DOMAIN 6 IS
 THE MOST REGULAR AND MAY BIND KIF DIRECTLY BY IONIC INTERACTIONS.
 CC DOMAINS 5 AND 7 ARE LESS WELL ORGANIZED AND MAY INDUCE FOLDS IN
 THE MOLECULE. DOMAIN 9 CONTAINS THE C-TERMINUS, CONSERVED AMONG
 DIFFERENT SPECIES.
 CC -1- PPM: KNOWN SUBSTRATE OF TRANSGLUTAMINASE. SOME 200 ARGININES ARE
 PROBABLY CONVERTED TO CITRULLINES BY PEPTIDYLARGININE DEIMINASE.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE S-100
 FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; Z18361; CA79165.1; -;
 DR EMBL; X51695; CA35992.1; -;
 DR PIR; A34209; A34209.
 DR PIR; S32633; S32633.
 DR PIR; A40691; A40691.
 DR HSSP; P02633; 31CB.
 DR InterPro; IPR001751; -;
 DR InterPro; IPR002048; -;

DR Pfam; PF01023; S_100; 1.
 DR Pfam; PF00036; efhand; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; FALSE_NEG.
 DR Repeat; Calcium-binding.
 KW DOMAIN 1 91
 FT CA_BIND 22 33 SITE I (LOW AFFINITY) (POTENTIAL).
 FT CA_BIND 62 73 SITE II (HIGH AFFINITY) (POTENTIAL).
 FT DOMAIN 413 832 14 X 28 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 413 448 1-1.
 FT REPEAT 449 476 1-2.
 FT REPEAT 477 504 1-3.
 FT REPEAT 505 532 1-4.
 FT REPEAT 533 560 1-5.
 FT REPEAT 561 588 1-6.
 FT REPEAT 589 616 1-7.
 FT REPEAT 617 648 1-8.
 FT REPEAT 649 678 1-9.
 FT REPEAT 679 706 1-10.
 FT REPEAT 707 742 1-11.
 FT REPEAT 743 771 1-12.
 FT REPEAT 772 796 1-13.
 FT REPEAT 797 832 1-14.
 FT DOMAIN 938 1507 23 X 23 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 938 961 2-1.
 FT REPEAT 962 985 2-2.
 FT REPEAT 986 1021 2-3.
 FT REPEAT 1022 1044 2-4.
 FT REPEAT 1045 1067 2-5.
 FT REPEAT 1068 1090 2-6.
 FT REPEAT 1091 1121 2-7.
 FT REPEAT 1122 1144 2-8.
 FT REPEAT 1145 1167 2-9.
 FT REPEAT 1168 1197 2-10.
 FT REPEAT 1198 1227 2-11.
 FT REPEAT 1228 1250 2-12.
 FT REPEAT 1251 1273 2-13.
 FT REPEAT 1274 1296 2-14.
 FT REPEAT 1297 1319 2-15.
 FT REPEAT 1320 1342 2-16.
 FT REPEAT 1343 1368 2-17.
 FT REPEAT 1369 1391 2-18.
 FT REPEAT 1392 1416 2-19.
 FT REPEAT 1417 1439 2-20.
 FT REPEAT 1440 1461 2-21.
 FT REPEAT 1462 1484 2-22.
 FT REPEAT 1485 1507 2-23.
 FT VARIANT 1145 1197 MISSING (IN SHORT FORM).
 FT VARIANT 1251 1273 MISSING (IN SHORT FORM).
 FT CONFLICT 1399 1399 E -> G (IN REF. 2).
 SO SEQUENCE 1549 AA; 201173 MW; E72FB9FF1326554E CRC64;

Query Match 2.98; Score 241; DB 1; Length 1549;
 Best local similarity 17.7%; Pred. No. 0.00034;
 Matches 229; Conservativity 185; Mismatches 518; Indels 362; Gaps 43;

OY 32 LTEERKIIIAVMDRQKKEEKESVLKIREHKAQP-----TQMFPSGITE 79
 DB 252 LLEERR-----EQLRRKEDRRDOQLRQDEATQEEISRGESRTSRQW-----Q 300
 OY 80 LVNVVLQPOOK---QPNKEPOTKLHQEFEMVKMGESQOQOBKGDAPTGCIGH 136
 DB 301 LESAADARQKRVYSRPHRQDSKRQEBELLEROEOQISEFVSLQDQGR-----Q 353
 OY 137 KTKFADCGHNCYSY-----CQTKFCARGGVSLSNKNVMVY---CNLCRQQETL 184
 DB 354 RLKQGRYDQWMPQLLEESQRRRTLYAKPAQROVVEEQLRLKEKRLQREKRRORE 413
 OY 185 TKSQAWFYNSGNTLQPDQKVPVPGRLRNEAPQEK-----KAKLHQDPQDAGDLS- 237
 DB 414 RQYREVELQREERLQREERQLOREERERKRRROREKQYLEKVELMEERQLOREERERRR 473

```

OY 238 -----VPAVEKGRAGLTPRODTIKNGSSGVKHQIASDMPDRKSPSPVSDONRREYOS 290
DB 474 QERERKQYLEKVELOEREOLOERERKROERKQYLEKVELOEREOLOERERKROER 531
OY 291 EEREDYQOYVPSDQTPMPSRPSDYADRSQRPQVE-----EPGLNTRDSNRGRHSHK 345
DB 532 KERO-YLEKVELOEREOLOERERKROERKQYLEKVELOEREOLOERERKROERK 590
OY 346 EYIVDEDEVSERDEROREREEYQARTRSDPNLARYPVKPPQVEQOMTHAEVSARHE 405
DB 591 QYL---EKVELOEREOLOERERKROERKQYLEKVELOEREOLOERERKROER 645
OY 406 RRHSDVSLANAE--LEDSRISILRMDRPSR-----QSVSEERRA 443
DB 646 KQYLEKELQERERQERQEE--QLTRERERKROERERQYLEKVELOEREOLOERERKR-R 703
OY 444 MENORSYMERTRERAGOSSTYPORTSNHSPPTPRSPITLDRPMRRAOSLRKHOLD 503
DB 704 QEREROY-LEK-EELQERERQERQEE-----LQERERKROERKQYLEK 750
OY 504 SAVRTRKREKMETLRNDSLSDSQSVPRPPPRHKKSGKMRQVSLSSSEELASTP 563
DB 751 LQOEDRLQERQERQEE-----REKR----- 772
OY 564 EYTSQDVELESESVSEKGSQKRRKRTSEQVGLSDSN-----TREKQKRMVYG 614
DB 773 QYLEKVELOEREOLOERERKROERERQYLEKVELOEREOLOERERKRRR--- 829
OY 615 GHSLEDELEWSEPOIKDSQVDTCCSTTLNESHSHKHPYTWQPSKDGDLIGRLILNK 673
DB 830 -QELRQERQEEELQERQDR-----KQFRDDOH---QNEVANSVYSHREKKE 874
OY 674 --RIKDGSPRDSGAMIGLVGKMGTEGRLCAFTTKVKGSLADTVGHILRDEVLW 731
DB 875 KSRQOLDMSVRESQFOQDLRLQODEOEK-----REBOEW 910
OY 732 NGRLLOGATPEEYVNIITLESKPPQYELVYSPRIGDIPRIPTSHADLESSSESOKM 791
DB 911 RSRQKRDQSPFA-----EQLLEEQOKETPRR 937
OY 792 DRHSIVTSPMSGMLRDVPOFLSGQLSIKLMFDKVGHOLITVTLAKADLPREDGPRN 851
DB 938 DR-----KREERQOLKQERKIRY-----LEDRKFR 967
OY 852 PVYKIYLPDRSDKNKRRTKTKTEPPKMNQFTTYSVHRRER-----ERMLITL 904
DB 968 EEOQLRLREEOQLRQERDRKFRRELSROERD-----RKFRERQOLQEREOQLR 1018
OY 905 WQDARRESESEFLGELLIEL-----ETALLDEPHWYKLQTHDVSSLPFRPS 953
DB 1019 QERDRKFREREOQLQEREREOQLRQERDRKFREREOQLQEREOQLRQERD----- 1068
OY 954 PLYPRROHGESPTRLQSRKRISSDSEVSDYDCEDGVGVSDYRNGDLOSTISVPO 1013
DB 1069 ----RKFREREOQLRLREEOQLRQ-----ENKFREREOQLRERE 1106
OY 1014 VMSNHGSPSSPRIRVYIGTRBSMSPSPAPQARNVEOGHCRATGHNNTISRMDRHY 1073
DB 1107 QLRLOEGEPOLRQK-----RDRKFHEEOLOEREO-----LRQERDRK 1147
OY 1074 MDDVSSDRDDBCAADROPYHRSNSTBORPLLERTTTRSSSE-----RPDTN 1122
DB 1148 FREBAQILKEHE-BQLRNGQERDRKFREREOLOERERELRQEREOLOERDRKFREREO 1206
OY 1123 LMRSMPSLMTGSRAPSPALSSHPRTGSVQVSPSTPTGRRGRLPOLPPKGLTERSA 1182
DB 1207 LLOERERK-----RQERERQOLRQERDRKFREREOQLQ-----EREOQLRQERDRKFRER 1258
OY 1183 MDIERENROMKLNKYKOVASDPRLEODYHNSKYR 1216
DB 1259 QLOERERQOLRQERDRKFREREOQLQEREOQLR 1292

```

```

RESULT 15
ATTR_MOUSE ID ATTR_MOUSE STANDARD: PRT: 2476 AA.
AC 061687;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TRANSCRIPTIONAL REGULATOR ATRX (X-LINKED NUCLEAR PROTEIN)
DE (HETEROCHROMATIN PROTEIN 2) (HPI ALPHA-INTERACTING PROTEIN) (HPI-BP38
DE PROTEIN).
DE ATRX OR XNP OR HP1BP2.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98213653; PubMed=9545503;
RA Picketts D.J., Taatzen A.O., Higgs D.R., Gibbons R.J.
RT "Comparison of the human and murine Atrx gene identifies highly
RT conserved, functionally important domains."
Mamm. Genome 9:400-403(1998).
RN [2]
RP SEQUENCE OF 325-1176 FROM N.A.
RX MEDLINE=97133299; PubMed=8978696;
RA le Douarin B., Nielsen A.L., Garnier J.-M., Ichinose H.,
RA Jeanmougin F., Losson R., Chambon P.
RT "A possible involvement of Tf1 alpha and Tf1 beta in the epigenetic
RT control of transcription by nuclear receptors."
EMBO J. 15:6701-6715(1996)
RN [3]
RP GENE EXPRESSION BY AFFECTING CHROMATIN.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PHO-FINGER DOMAIN.
CC
CC This SNF2-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AF026032; AAC08741.1;
CC DR EMBL: X99643; CAA67962.1;
CC DR MGI:103067; Xnp.
CC DR InterPro: IPR000330;
CC DR InterPro: IPR001650;
CC DR Pfam: PF00176; SNF2_N.1
CC DR DNA repair; Nuclear protein; DNA-binding; Helicase; ATP-binding.
CC KW DOMAIN 219 267
CC FT NP BIND 1579 1586
CC FT SITE 1704 1707
CC FT DOMAIN 319 322
CC FT DOMAIN 735 738
CC FT DOMAIN 1001 1004
CC FT DOMAIN 1130 1135
CC FT DOMAIN 1182 1185
CC FT DOMAIN 1238 1245
CC FT DOMAIN 1484 1487
CC FT DOMAIN 1924 1931
CC FT DOMAIN 2205 2208
CC FT DOMAIN 2245 2248
CC FT DOMAIN 2403 2408
CC FT SEQUENCE 2476 AA; 278601 MW; 90A42B790FC4FF4C CRC64;

```

Query Match 2.9%; Score 239.5; DB 1; Length 2476;
 Best Local Similarity 19.0%; Pred. No. 0.00071;
 Matches 329; Conservative 231; Mismatches 616; Indels 553; Gaps 80;

```

QY 37 KRIIAVNDROKKEEKEOSVLKIKEHKAOPTOMF-----PFGITELVANNVLO--P 87
Db 240 KICILRNIGR-----KELST--IMDEN-----NOMYICIQPEPLLDIVLACNSVFENLE 287
QY 88 OOKQONER-----EPQTLHQO-----FEMTKEOVKKMGESOOQOQOKG----- 127
Db 288 QILOONKRRKIKVDSERTKSVCDQTSKFSFKSSSCNGEKKLEESCSCGVSTYSHAL 347
QY 128 DAPTCGICHTKFCFADGCGH--NCSYCOTKFCARGCGRVSLRSNNVMWVNCN--CRKQOEL 184
Db 348 SVPKEMIKTKTTLIETSMNSY--IKFLQOANDSEKTS--AMALCQLSKTSKSVLDDI 403
QY 185 TKSQAMFYNSGNTLQOQPD-----QKVPGLRNEAPQO-----KKAKLEQPOFOGAP 234
Db 404 KKAHLAEEDLSEIOALDDVHKKEKNTKDLKSTDAKSETKLGEKESYSTEREPLKIDA 463
QY 235 DLSVAV--EKRAGLTQODITKNGSVKHOIASDMPDRKSPSVSRQONRYESEE 292
Db 464 RSVKAIQDEQERAKHSTSGE--HKSGS-----RKDGSOYEPTNT 501
QY 293 REDYS-----QYVPSD-----GTMPSPSDY--ADRSOREPOFEYEPGHLNRYDSN 337
Db 502 PEDLMDIYVSVSSVPEDIFDLSLEAMEVQSSADYOGDNGSTEPELESSESVKLNYSKD 561
QY 338 RGHHRHKEYIVDDEVEDERDEYERQOREEYQARYSD-----PNLARYPVKPOYE 390
Db 562 SGNKJRSK-----VYAKVKELFVKLTPLVSLNSPIKGVDCQ 598
QY 391 EOMRHAHYEYRAHERRHSDVSIANALEDSTRSLRMDRPSQORSYSEKRAAMENORST 450
Db 599 -----EVSOKNGRKSSGVA-----RSSEKCPREELISHENNVTI 634
QY 451 SMERTREAOGOSYPORTSNHSPPTPRSPPIPLDRPDMRADSLRQOHLDPSSAVAKTK 510
Db 635 LLEDS-----DLRSPRYKTTPLRQOTESNPMASNDDESNGMK 675
QY 511 REKMETMLRNDLSLSDQESVRRPPRPKSKGKGMROYST-----SSSEELASTEPE-- 564
Db 676 KQKMGPIRKK--DKRNSADCATDNPKPKHVPKA--KQVIGDQNSDDEMLAVLKEAS 730
QY 565 ---YSCDVELESESVSKG-----DSQKGRKTSQCVLNSNTRS-----ERQ 607
Db 731 ONGHSSSTDTINEPQMNHKGTKGDKDNGKRKNKSTSG--SDFTKKKGSTETSIISK 788
QY 608 KKRATYGGHSLSEEDLEWSEPOIKDSGVDTCSSPTTLNEHSHSDKHPTWQPSKDGRLI-- 666
Db 789 KRONTSSESNYSSELEREIKTMSRIGAKRSYPERKEDESSDEB-----KQKGVVD 840
QY 667 -GRILLNRKLDGVPYRDSGAMLGKVGKMTESGRLCAFITRKVKGLADTVGHLRPG 725
Db 841 NGHERAKTTJGSSADDTGDTGRQ--GGSCSIAG--GSIKVRSGV-----EFR-- 887
QY 726 DEVLENGNLLQOGATEEYVNTILESKPEPOVELVYSRPIGDIPIRIPDST----- 775
Db 888 -EMLCKPGVSSDGA-----EKPSYKEENVNSPEDKRVSKTKKTKHLRSRQSRK 935
QY 776 -----HAOLESSSSSFESOKMDRPSISVTSPMSPGMLRDVQFLSGOLSIKLF 824
Db 936 GKGSSDGTDRPPKPKQSDSSSEGEKKQ-----SKRPGTGKAKAPDLGE----- 981
QY 825 DKVGHQIIVTILGAKDLPSREDGRPRNPYVKIYFLPDRSD-----KNKRTKTVKKTLEP 879
Db 982 -----TLKREQEMDSSDQTER-----LPEEEHIGPFSKGIKQSKTDITAG 1024
QY 880 KKNQOTFIYSPVHREPREMLETITLMDQARVREBESEFLGELLIELETA--LLDDEPHWK 938
Db 1025 K-----GKKWKDKSCF-----KKEELSDVDKLPKGKSDSCSSSEDKTNR 1065
QY 939 LQTHVSSLLPLRPSBYL-----PRQLHGESPTRLQRSKRIDSSEVSDYDCEGV 990
Db 1066 VSLREKKRSLPAKSPGRPECCSSDTEKSLKQCCDSTEKRRKRIDLRERN----- 1118
QY 991 GYVSDYRHNGRDLQSTSLVPEQVMSNHCSFSGSPHRYVIGRTNSWSPAPPORAVE 1050

```

```

Db 1119 ---SSSKRNTKEVKSAS-----SSSDAEGSSEDNKKOKKQRT-----SAKKKTGNK 1162
QY 1051 QGHRGTRATGIVNTISNRDRHVRVMDHYSDBDRDCEADRPYHRSRTEQ--RPLER 1108
Db 1163 EKKR-----NSLRATPKRKQVDTSSS--DIGDDONSAGEESSDECKIKPVEN 1211
QY 1109 TPTRSRS--SERPDNLRSMPSLMTGRSAPPSP-----ALSRSHPTGVS 1152
Db 1212 LVPLSHHGFQCGSSGDEALSKSVATVDDDDNDNDENRIAKMKLLEIKANLSDEBSS 1271
QY 1153 QTSPPSTPGTGRGRQRLPOLPPKGLTERSAMDIEERNQOMKINKYQVAGSPRLBODYH 1212
Db 1272 DDEPDG--GKKRRIQOSESP-----ADGELRRQDLAVQVNSSESDSDS--EESK 1320
QY 1213 SKYSGMDPH-----RGADYVSTSSDSDVDY--SAVSRTS 1247
Db 1321 PRYHRLIRHLTLUSDGSEGEKPTKPEHKEAKRRNRKVSSEDESDTDFQESGVSEV 1380
QY 1248 SASRFSSTSYMSVOSEPRGRNKRISV-----FTSKMONRQMGVSGKNLTGSTSISGDMC 1301
Db 1381 SESE-----DEQRPRTRSKAKKALEENORSYKQKKRRRIKQVEDSSSEKSHSE-- 1431
QY 1302 SLEKNDGSDPTAVGALGTSGKRRRSSIGAKVVALVGLSRKSRSASQLSQTEG--GKKLR 1360
Db 1432 --DKKEGDEEDE-----DEDEDEDEDENDDSKSPGKGRKIR 1466
QY 1361 STVQSTFETGLAVEMRNMTQASRESTDGSNYSSEBGNLFPVGRLASDQSFDFDG 1420
Db 1467 KILK--DOKLRTETONALKEEERRK-----RIARERERERK-- 1502
QY 1421 LGPAOLVGRQTL--ATPAMGDIQVM--MDKGOLEVEIIRA--RGIYVKGPSKTLPAVY 1474
Db 1503 -----REVIEIEDASPTPCITTKLYLDENETKEPLVGVHNMVTK----- 1544
QY 1475 VKVYLLDNGV-----CIAKKTIVARKTLEPLYQQLSFEESP-----QGRV 1516
Db 1545 LKPHQVD--GVQFMMDCCCESEVEKTK-----KSPGSGCIIAHCMGLGKT 1586
QY 1547 LQIIVMGDYGMDHKSFMGVAQIILDELELS-----NVIYGNPKLF 1557
Db 1587 LQVV-----SFLHTV--LLCDKLDFTALVYVCPNLALNMWNEF 1623

```

Search completed: November 21, 2001, 16:10:18
 Job time: 239 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:00:24 ; Search time 19.55 Seconds
(without alignments)
1830.191 Million cell updates/sec

Title: US-09-617-099b-1

Perfect score: 8285

Sequence: 1 MSAPLGPGRAPTRPAASQP.....TRRASQSSLESNGPYSRS 1590

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCrnUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	3.7	1898	1 US-08-056-200-94	Sequence 94, Appl
2	307	3.7	1898	2 US-08-800-644-94	Sequence 94, Appl
3	258.5	3.1	2842	1 US-07-741-940-7	Sequence 7, Appl
4	258.5	3.1	2842	1 US-08-289-548A-7	Sequence 7, Appl
5	258.5	3.1	2842	1 US-08-452-654-7	Sequence 7, Appl
6	258.5	3.1	2843	1 US-08-452-655B-2	Sequence 2, Appl
7	258.5	3.1	2843	1 US-08-450-582-2	Sequence 2, Appl
8	258.5	3.1	2843	3 US-08-450-582-2	Sequence 2, Appl
9	258.5	3.1	2843	3 US-08-450-582-2	Sequence 2, Appl
10	258.5	3.1	2973	2 US-08-821-355A-7	Sequence 7, Appl
11	258.5	3.1	2973	2 US-09-003-687A-7	Sequence 7, Appl
12	258.5	3.1	2973	2 US-09-136-605-7	Sequence 7, Appl
13	250.5	3.0	2843	1 US-07-741-940-2	Sequence 2, Appl
14	250.5	3.0	2843	1 US-08-289-548A-2	Sequence 2, Appl
15	250.5	3.0	2843	1 US-08-452-654-2	Sequence 2, Appl
16	250.5	3.0	2843	1 US-08-452-654-2	Sequence 2, Appl
17	225.5	2.7	1780	1 US-08-370-235A-5	Sequence 2, Appl
18	225.5	2.7	1780	1 US-08-769-309A-5	Sequence 5, Appl
19	222	2.7	2101	1 US-08-994-570-5	Sequence 5, Appl
20	222	2.7	2101	1 US-08-466-390-4	Sequence 4, Appl
21	222	2.7	2101	1 US-08-467-781-4	Sequence 4, Appl
22	222	2.7	2101	1 US-08-470-950-4	Sequence 4, Appl
23	222	2.7	2101	2 US-08-483-924-4	Sequence 4, Appl
24	219	2.6	2101	1 US-09-452-294-1	Sequence 1, Appl
25	219	2.6	2101	1 US-08-195-487-4	Sequence 4, Appl
26	206	2.5	1805	5 PCT-US93-06160-4	Sequence 2, Appl
27	205	2.5	1162	2 US-07-853-913-2	Sequence 2, Appl
				2 US-08-728-323A-2	Sequence 2, Appl

28	200	2.4	927	3 US-08-895-601-6	Sequence 6, Appl
29	197.5	2.4	3669	4 US-08-061-376-5	Sequence 5, Appl
30	195	2.4	1068	1 US-08-396-479B-12	Sequence 12, Appl
31	195	2.4	1068	1 US-08-818-823-12	Sequence 12, Appl
32	193.5	2.3	1658	2 US-08-609-049A-13	Sequence 13, Appl
33	193.5	2.3	1658	4 US-09-170-996-13	Sequence 13, Appl
34	193.5	2.3	1726	2 US-08-609-049A-30	Sequence 30, Appl
35	193.5	2.3	1726	4 US-09-170-996-30	Sequence 30, Appl
36	190	2.3	1360	4 US-09-393-569-2	Sequence 2, Appl
37	188.5	2.3	1075	5 PCT-US94-07297-41	Sequence 41, Appl
38	185.5	2.2	1210	1 US-08-320-559-26	Sequence 26, Appl
39	185.5	2.2	1210	3 US-08-545-860D-26	Sequence 26, Appl
40	185.5	2.2	1210	5 PCT-US94-04496-26	Sequence 26, Appl
41	182.5	2.2	137	2 US-08-609-049A-15	Sequence 15, Appl
42	182.5	2.2	137	4 US-09-170-996-15	Sequence 15, Appl
43	182	2.2	1706	2 US-08-459-568-2	Sequence 2, Appl
44	182	2.2	1706	2 US-08-399-411-2	Sequence 2, Appl
45	182	2.2	1706	3 US-08-516-859A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-056-200-94
Sequence 94, Application US/08056200
Patent No. 5616500
GENERAL INFORMATION:
APPLICANT: Steinert, Peter M.
APPLICANT: Lee, Seung-Chul
APPLICANT: Kim, In-Gyu
APPLICANT: Chung, Soo-Il
APPLICANT: Park, Sang-Chul
TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056,200
FILING DATE: 30-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fedrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 1898 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-056-200-94

Query Match 3.7%; Score 307; DB 1; Length 1898;
Best Local Similarity 17.4%; Pred. No. 2.8e-14;
Matches 245; Conservative 213; Mismatches 514; Indels 434; Gaps 53;
QY 19 QPPPEPMPPLSHUTEE--RKILAVMDRQKREKQSVLKIKERKAPPTQWFPSPG 76

```

Db 264 EPQORELOE-----EEQORLKEOELRREOEEQOQORLREOQLRKEOE----- 312
QY 77 ITELANNVLOPOOKORNEKEPOTKILHOQFMAYEOKVAKKEESOOO---EOKGADPTCG 133
Db 313 -----ERRQOEERREOQORREOQERREOQLRREOEEERREOQLRREOE----- 358
QY 134 ICHKTAFADGCGHNSYCQKFCARCGRVSLSNKVMWVCNLCKOQELLTKSGAMFYN 193
Db 359 -----RREOQL----- 364
QY 194 SSSNTLOQPDQKVPRLRNEAPQOEKAKLHEQPOFQAGBLSVPAVEKRAHGLTRQD 253
Db 365 -----RREOEEERREOQLRREOEEERREOQLRREOQLRREO-----QJLREOQLRREQ 412
QY 254 TIKNSGVNHOJASMDPSDKRSPSPSRQONKRYEGSEER-EDYSQVPSDGTWPSPSD 312
Db 413 QJLREOQLRREOQLRREOQLRREOQLRREOQLRREOEEERHEQKHEOERREOQLRKEOE 472
QY 313 YAD-----RRSOREPOFYEEEPGH---LNYRDSNRGRHSHSEYIYDDEVE 355
Db 473 RRDWLKREETERHEOERKQOLKROEERRRRWLKEEERREOQERREOQLRKEOE 532
QY 356 SPOE-YEROREEEXQARYRSDPNLARYPVQPYEOMKRIAEVSRARERHSDVSLA 414
Db 533 RREOQLRKEOEEERLOQRLREOQLR---EOEERLEQLKREERKLEOERREOQLR--- 586
QY 415 NAELEDSRISLMDRPSQORSSEKRAMENORSTSMETRAQOSSIPQTSNHSPP 474
Db 587 KROEERROQLKREERROQLKREOERLEOR-LKREVERLEOERDERLKRKEPE 645
QY 475 YRRSPIDPRDMRADSILRKOHLDPSAAVAKTKREKMETLRNDSLSSDQSESVRP 534
Db 646 EERRH-----ELKSEOEERH---EQLRREOQERREOQLRKE---EEERLEOR 690
QY 535 PPRPKSKKGGMMROYSLSSSEBELA-----STPYTSCDDVELESSEVSEKGDQOKR 588
Db 691 LKREHEER---RQOELAEEOQARERIKSRIPKW---QWOLESEA---DARQSK 737
QY 589 -RKTSEOGVLSDSNTRSEKOKRMVYCGHSLDELDWSEPOIKDSQVDCSSTLMEHS 647
Db 738 VILLEPQAGRAAPQOEKKRR---BSELQOE-----EERA 772
QY 648 H-----SDKHEVTWOPSKGDRLIGRLLNRLKDGSPRDSGAMIGLVGKMTESG 701
Db 773 HROOQEEERQRPFTWQOEKESRGORLSAR---PLRQORQRLAERROQREQR 827
QY 702 RUCATITVKKGSLADYVGHLPDGVLEWNGRLLQAGATFEVYNTILSKPEPOVELV 761
Db 828 FL-----PEEEKEORGR-----QRREKELOF 851
QY 762 SRPIDIPRIDPSTHAOLESSSSSFESOKMDRPSISVTSMPSPGMLRDVPOLISGOLSK 821
Db 852 -----LEE-----EEOLOREKRAQOLOEEDDELOEDERRROQERQD 889
QY 822 LWFQVGHQILVTLIGAKDLPSEDEGRPNPYKTYFLPDRSDKNKRTKYAKTKLEPKW 881
Db 890 KW-----RQWL-----EERKRRRHT-LYAKPALQOLKKEOQLLOEEEBE 931
QY 882 NOTFIYSVHREPEREMLLEITLMQARVRESESEFLGILILETALIDDEPHWY---- 937
Db 932 REE---RKRROEOEROYR---EEOLOQOEEOQLRE---EREKRRROEROEROYRKDK 981
QY 938 KLOTHIVASLPLRPSPLPRROLGESP--TRLORSKRISDSEVSDVDCEDGVGVSD 995
Db 982 KLOQKE-----EQLGEEPERKRRROEREKTYREE----- 1011
QY 996 YRHNGRDLOSSITSLVEQVMSNNHSCSPHARDVIGTRTSMW---PSAPPORVNEOG 1052
Db 1012 -----ELOE-----EEOQLRERREK-----RRQWMEYOYRKDKDLQOEEDL 1050
QY 1053 HRCGRATGHYNTISMDBRHVDDHYSSDRDRCADARQYHRSRSTEQRLLETTR 1112

```

```

Db 1051 LREER-----EKRLROERROYREBELOEEOQLGEEERETRRROELEROYRK 1099
QY 1113 SRSSRPDTNMRSPSLMTGRSAPSPALSKRSHRTGSVQTSPTSSTGTGRGQOLPOL 1172
Db 1100 EEELOQEEEOQLRE-----EPKRRROERROCREEBEELQOEEOQLMEERKE 1147
QY 1173 PPKGTLEASMDIERNNOMKLINKQVAGSDPRLEQDYHSHKYSRGWDPH----- 1222
Db 1148 RRROELEROYEEELQOKKRRQYR-----DEQORSDLKQWQEPKEKENAVRDNKV 1198
QY 1223 --RGADTVSTKS-SDSDVSDVSAVSRTSSASRFSSTYSWYOSEPRGRNKISVFTSKMO 1279
Db 1199 YCKRENOQFQOLDSQYRD-----KQSOQDLOHLLGEOQERDR-----EOE 1240
QY 1280 NRQNGVSGKNLTSTSTSGDMCSLEKNDGSDQTAVALGTSGKKRRSSIGAKVAYIVGL 1339
Db 1241 RRMQOANRHPREE-----QLEREBQKA-----KRR----- 1268
QY 1340 SRKRSASQLSQTEGGCKKLASTYQR 1365
Db 1269 DRKSQOEKQLRREERERRRROETDR 1294

```

```

RESULT 2
US-08-800-644-94
; Sequence 94, Application US/08800644
; Patent No. 5938752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054,001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ. ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-800-644-94

```

Query Match 3.7%; Score 307; DB 2; Length 1898;
Best Local Similarity 17.4%; Pred. No. 2,8e-14;

MOLECULE TYPE: protein
 ORIGINAL SOURCE: ORGANISM: Homo sapiens
 IMMEDIATE SOURCE: CLONE: APC
 US-07-741-940-7

Query Match 3.1%; Score 258.5; DB 1; Length 2842;
 Best Local Similarity 17.9%; Pred. No. 2.7e-10;
 Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69;

30 SHLSEEEKILLAVMDRQKKKEEKEQSVLKIKEEHKAPDTQMPPEPGCITELVNNVLPQO 89
 1011 NHMDNDDELTPPIYSLKYSDEQ-----LNSGRQSPQNERAARPHITE--DEIQSEQ 1064
 QY 90 KQPEKEPQTLHQEFYMEKQVKMGESQO---QOEQKGAPTGICHTKTFADCGH 146
 Db 1065 RQSRNQ-----TTPYVTESTDKHLKFPHPGOE-----CVSPYRSKANGSET 1111
 QY 147 NCSTCQTKFCARCGRVSIRSNKVMVCMLCRKQDEILTKSGAMFYNSGNTLQDPQKY 206
 Db 1112 N-----RVGKNHGINQNVQSILC-----QEDDYEDDKPTNYSERYSEEHQHEE-- 1154
 QY 207 PGLRNEAPQEKKAKLHEQPOFGAPGDLSPVAVKGRAGHGLTRDITKNSGVKHQIA 266
 Db 1155 -----EERPIYSLIKYNEKRRHVDQPIDYSL-----KTA 1183
 QY 267 SDMSDRRSPSVSRDQNKRYEQRSEREDYSQVPSDGTMPSPDYADRRSQREPFYE 326
 Db 1184 TDIPSSQKQSFSSKSSSGQSSKTE-----HMSSENTSTPSSNAKRONLHSSAQ 1236
 QY 327 EPGHLNYSNRGRHRSKE---YIVD-----EDYE 335
 Db 1237 SRSQGPQKAACCKVSSINQETIQTYCVEDTPICFKSCSLSSLSAEDIGCNOTQEAD 1296
 QY 356 SRDEXERQREERYQARYRSDPNLARYPKPQPYEFQMRTHAEVSRARHERHSDVSLAN 415
 Db 1297 SANLIQLEIKETIKTRSAEDVEYPAVSQHPRTKSSRLQGS--SLSSSARAKAVEFSS 1355
 QY 416 AELEDSRISLMDRP-----SRQSVSERRAMNQSVSMERTREAO---- 459
 Db 1356 GAKSPKSGAQPKSPPEHYVQETPLMFSCITSVS---SLDSFEERSIASSVQSEPCSG 1411
 QY 460 -----GOSTPQRTSNISPPTRPSRPIPLDRPMKRADSLRKHQHLDPSSAVAKTKRREK 513
 Db 1412 MVSGLIISPSDLPDPGQTMPPSPRSKTPPP-----PQTA--QTKRE- 1450
 QY 514 METMLRNDLSLSDQSESVRPPRPKSKKGGKMRQVSLSSSEEL-----ASTPEYTSQ 568
 Db 1451 ---VFKKAPTAERKES-----GPKQAQVNAVAQRYVLPADATLLHFATESTPDGFCSC 1501
 QY 569 D-----DYEL-----ESESVESEKGSQ--KGRKRT--SEQ 595
 Db 1502 SSSLSALSLDEPFIOKQVQLRIMPVQENDNGNETESEQPKESNENQEKAEKTIIDSEKD 1561
 QY 596 VLSDSN-----TRSEROKKRMVYGGHSLSEDLWSEPDITKDSGVDTCS 639
 Db 1562 LLDDEDDDDIEILLECIIISAMPTKSSRKAKKPAQATASKLPPVARKRPSQ-----PVYK 1615
 QY 640 TTLNEHSHSKDHPVTWQPSKXGDRL--IGRIILL-----KRLDGSVPRDSGAMGL 690
 Db 1616 LIPSONRLQPKH--VSFTPGDMPRYVCEGTPINSTATSLDLTIESPPNLAGEGV 1674
 QY 691 KVVGGKATSEGRICAFITKVKKSLADIVGHLRPGDEVLEWNGRLIQAGTFEEVYNITLE 750
 Db 1675 R--GG--AQSG-----EFERKDTIPTEG--RSTDEA-----QGCTSSVTIPEILD 1713
 QY 751 SKPEQVELV-----SRPIGD-----IPRIPDSTHAOLESSSSSFESOKMDRPSISVT 799
 Db 1714 DNKAEEGILAEINCISAMPKGSKHPFRVKKIMDQVQ--QASASSSPANKQLDGGKKKKPT 1772
 QY 800 SPMSPGMLRDVQPLSGQLSIKLMFDKVGHQLIVTILGAKDLPSREDGRNRNPYVATYFL 859

Db 1773 SPVKP-----IPQNTERTRY-----RKNDSKNNLNAAERVF 1804
 QY 860 PDRSDKRRRTKTVKKTLEPK-----WNQTFYSPVH-----RR 893
 Db 1805 SDNKSCKQNLKNNKSKDFNDKLPNNEDRVRGSAFDSPPHATYPIEGTPYCFSRNDSLSSL 1864
 QY 894 EFRERMLEITLMQARRESESEFLGELLIELEFALLDDERHWKIQYHDVSSULPR-- 951
 Db 1865 DFDDDDVDLNR-EKAEILKKAENKRESEKAVYSHETLSNQOSANKTQA--TAKQFINRGQ 1921
 QY 952 PPSYLPRL-----HGESPTRLQR-----SKRISDEVSDDYCE----- 987
 Db 1922 PKPILQKSTFPQSSKDLIPDGAATDEKQNFATIENTPVCFSHNSLSSLSIDIDENNKK 1981
 QY 988 -----DGVGVSDYRHNG-----RDLQSTLSVPBQVSSNHC 1020
 Db 1982 ENERIKETEPPDSQGEPSKPAQASVAPKSFHVEDTPVCFSSNSSLSSLSIDEDLLQEC 2041
 QY 1021 SPSSPHRVD---VIGTRSMSPSAPPQRVNVEQGHGRTATGHNTISRD--RHRVMD 1076
 Db 2042 ISSAMPKKKKRSLKGNKHSKSP-----RNM--GGILGEDTLTKDQRPDSEHGLSPD 2094
 QY 1077 HYSDDRDRDCEAD-----RQ-----PYHRSRSTE 1101
 Db 2095 SENFDKALQEGANSIYSSLHQAANAACLSRQASDSDSLILSLKSGISLGSFPHITPDE 2154
 QY 1102 QRPL-----ERTTSSRSERPDIN----- 1122
 Db 2155 EKPTSNKGPRIKPKGKSTLETKEESKIGKGVKSLITGKVRNSELISQMKQ 2214
 QY 1123 -LMSMSLSMTGRAPSPALSRSHPRGVS-----QTSPSSTPGGRGRQPLPOLP 1174
 Db 2215 PLOANMSPISGRPMIHPGRNSSSTSPVSKKPPILKTPASKSPSEG---QIATTPSP 2270
 QY 1175 KGLTERAMDIIEERNOMKILNKYQVAGSDPRLQDYHSKYSRG---WDPHRGAD----- 1226
 Db 2271 RGAKPSYKSELSPARQT-----SQIGSS-----KAPRSGRSDTPSPRAQPLSR 2318
 QY 1227 TVSTKSSDSVDVSAVSRTSSASRFSSTYSWVOSERPGRNRIKISVFTSKMNQOMGV 1286
 Db 2319 PIQSPGRNRSIPGRNGISPPKRLQPLPTSPSTASTKSSSGSKMS--YTS--PGROM--S 2373
 QY 1287 GKNTKTSISGDMCSLEKNQSGSDTAVGALGTSKKRRBSIGAKMVAITGLSKKRSRA 1346
 Db 2374 QQNTLKQGTGSKNASLIPRSASAKGLNMNNGANK-----VELSRSSFTK 2422
 QY 1347 SOLSQTEGGGKTL-----RSTVQRSTETGLAVENRNMNTROASREST----- 1388
 Db 2423 SSGSESDRSERPVLYRQSTFIKEAPPTLRKRLIESASFEELSPPSSRPASTPTRQAGQTPV 2482
 QY 1389 -----DGSNMSYSS 1397
 Db 2483 LSPSLPDMSLSTHSS 2497

RESULT 4
 US-08-289-548A-7
 ; Sequence 7, Application US/08289548A
 ; Patent No. 5648212
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTSEN, HANS
 ; APPLICANT: ANAND, RAKESH
 ; APPLICANT: CARLSON, MARY
 ; APPLICANT: GRODEN, JOANNA
 ; APPLICANT: HEDGE, PHILIP J.
 ; APPLICANT: JOSLYN, GEOFF
 ; APPLICANT: KINZIER, KENNETH
 ; APPLICANT: MARKHAM, ALEXANDER F.
 ; APPLICANT: NAKAMURA, YUSUKE
 ; APPLICANT: THLIVERIS, ANDREW
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
 ; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

```

:      NUMBER OF SEQUENCES: 102
:      CORRESPONDENCE ADDRESS:
:      ADDRESSEE: Banner & Allegretti, LTD
:      STREET: 1001 G Street, NW
:      CITY: Washington
:      STATE: D.C.
:      COUNTRY: USA
:      ZIP: 20001-4598
:      COMPUTER READABLE FORM:
:      MEDIUM TYPE: Floppy disk
:      COMPUTER: IBM PC compatible
:      OPERATING SYSTEM: PC-DOS/MS-DOS
:      SOFTWARE: Patentin Release #1.0, Version #1.25
:      CURRENT APPLICATION DATA:
:      APPLICATION NUMBER: US/08/289,548A
:      FILING DATE: 12-AUG-1994
:      CLASSIFICATION: 435
:      ATTORNEY/AGENT INFORMATION:
:      NAME: Kagan, Sarah A.
:      REGISTRATION NUMBER: 32,141
:      REFERENCE NUMBER: 1107,46943
:      TELECOMMUNICATION INFORMATION:
:      TELEPHONE: 202-508-9100
:      TELEFAX: 202-508-9299
:      INFORMATION FOR SEQ ID NO: 7:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 2842 amino acids
:      TYPE: amino acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      MOLECULE TYPE: protein
:      ORIGINAL SOURCE:
:      ORGANISM: Homo sapiens
:      IMMEDIATE SOURCE:
:      CLONE: APC
:      US-08-289-548A-7

Query Match      3 1%; Score 258.5; DB 1; Length 2842;
Best Local Similarity 17.9%; Pred. No. 2.7e-10;
Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69;

QY 30 SHLTTEERKIIIAVMDROKKEEKESVLKIKENHKAOPTOMPFPSGITELVNVNVLDPQ 89
DB 1011 NIMDNDGELDPINYSKLSDEQ----LNSGRSPSQNRMARPKHITF--DEIKSEQ 1064
QY 90 KQPKNEPQTKLHQFEMTKEOYKKMGESQ--QEOQKGAAPTCGICHTKFPADCGH 146
DB 1065 ROSRNO-----TTPYVTESTDDKHLKFPHFQGOE-----CVSPYRSRGANGSET 1111
QY 147 NCSYCTKFCACGCGVSLRSKKVMVVCILCRKOQELITKSGAMFYNGSSNTLQDPQKV 206
DB 1112 N-----HVGSNHGINQNVSSGLC-----QEDDYEDDKPTNYSERYSEEGHEE-- 1154
QY 207 PGLNNEAPQEKKAKLHDPQFOGAPCDLSVPVAVEKGRAHGLTROPTIKNGSGVKQIA 266
DB 1155 -----EERPTVYSIKYNEKKNHVDQPIDYSL-----KYA 1183
QY 267 SMPSDRKSPSVSRDQNRREYSEEREDYSQYVPSDGTMPRSPDYARRSQEPOFYE 326
DB 1184 TPIPSQOKGFSFSSKSSQSSKTE-----HMSSESENTSTPSSNAKRONOLHPSSAQ 1236
QY 327 EGHLYNRNSNRGRHNSK-----YIVDD-----EDVE 355
DB 1237 SSGGQPKAKATKCVSSINQETIQTICVEDTPICFSCSSLSLSSADETIGCNOTTDEAD 1296
QY 356 SRDEYERQREEEYQARYSDNPILARYPVKPPQYEQMRTHAEVSRARHRRHSDVLAN 415
DB 1297 SANTQIAIKIKIKIGTRSMDEPVSEVPASQHPRTKSRIGGS--SLSSSEARHKAIVFSS 1355
QY 416 ALEEDSRISLLMRDP-----SRQSVSERRAAMENQKSYSMERTREAO----- 459
DB 1356 GAKSPSKSGAQTTPKSPPEHYVQETPILMFSRCTSVS-----SLDSFESKSIASSVQSEPCSG 1411

```

```

QY 460 -----GSSYQRTSNHSPPTPRRSPIDLPDRDMRADSLRKHHLDPSSAVRKTREK 513
DB 1412 MVSGLISSDLPDSFGQTMPPSRKSTPPP-----PQOTA--QTKRE-- 1450
QY 514 METMLRNDLSLSDQSESVRPPPRPHKSKGKGMQVSLSSSEEL-----ASTPEYTCG 568
DB 1451 ---VPKNKAPTAERKES-----GPKQAAVNAAYORVQVLPADVTLIHPATESTPDGFGC 1501
QY 569 D-----DVEL-----ESESVSQKSGQ--GKKRKT---SEGG 595
DB 1502 SSSLSALSLEDPFLQKDLVLRIMPPVQENDNGNETSESEQPKSNENQEAETIDSEKD 1561
QY 596 VLSDSN-----TRSEROKRRMYYGSHSLEEDLEWSEPOIKDSGVDTCSG 639
DB 1562 LLDSDDDDDIIELEECIISAMPTKSSRKAKRAQASKLPPPVARRPSOL-----PVK 1615
QY 640 TTLNEHSHSKHPVYTWOPSKDGRLL--IGRIILN-----KRLKGSVPRDSGAMLG 690
DB 1616 LLPSONRLQPKH--VSFTPPGDDMPRVYCVGECTPINFSTATSLDLTIESPPNELAGEGV 1674
QY 691 KVGSGKMTESGRLCAFITKVKKGLADYVGHLPDDEVLEMMGRLLQGTFFEVYNIILE 750
DB 1675 R--GG--AQSG-----EFERDRTIPEG--RSTDEA-----QGGKTSVTTIPELD 1713
QY 751 SKPEQVELVY-----SRPIDG-----IPRIDSTHAQLESSESSFEESOKMDRPSISVT 799
DB 1714 DNKAEEGDIILAEICINAMPKGSHKPRYKKTIMDOYQ--QASASSAPKNNQDLGKKKKFT 1772
QY 800 SPMSPGMLRDVPOFLSGQLSTKLMFDKVGHQLIVITLAKDLPSREDGPRNRYKIYEL 859
DB 1773 SPVKP-----IPQNEYERTRV-----RKNADSKNNLNAERYF 1804
QY 860 PDRSPKMKRRTKYTKTEPK-----NQOTIYSPVH-----RR 893
DB 1805 SDNKSCKKQNLKNNKSDKNDKLPNNEDRYVGSFADSPHHYPIEGTPYCFPSKNDLSST 1864
QY 894 EFERMELITLMDQARVRESESEFLGEILILETALDLDEPMHYKQTDVSLPLPR-- 951
DB 1865 DFDDDDVDLIR--EKALRKAKENKSEAKVYHTTELTSMQOSANKTQA--IAKQPINRKQ 1921
QY 952 PSPYLPRROL-----HGSEPTRLQV-----SKRISDEVSVDYDCE--- 987
DB 1922 PKPILQKQSTPQSSKDLIPDGAATDEKLQNFALIENTPYCFSHNSLSLSDIDQNNMK 1981
QY 988 -----DGVGVSDYRNG-----RQLOSTLSVPEDVYSSNHC 1020
DB 1982 ENERIKETEPDQGEPSKPAQSGYAPKSFHVEDTPVCFSRNSSLSTLSDSEDLQEC 2041
QY 1021 SPGSPRHVND---VIGRTFRSPSPAPRQORNEQGHGRTARGHYNTISRMD--RHRVMD 1076
DB 2042 ISSAMPKPKKPRSLRGDNKHSF-----RNM--GGILGEDLILDLKDQRPDSHGLSPD 2094
QY 1077 HYSDDRDRDCEAAD-----RQ-----PYHRSRSTE 1101
DB 2095 SENFDWKAIOEGANSIVSLHQAAAAACLSRQASSDSDSILSLKSGISLCSPEHLTPDQE 2154
QY 1102 QRPLL-----ERTTSSRSSEPDN----- 1122
DB 2155 EKPTFSNKGPRILKDGKSTLETETKIEESKGIKGGKYYKSLITGKVSNSIEISGOMKO 2214
QY 1123 -LMRSPSLMIGRSAPSPALSRSHPRTGSV-----QTSPTSGTQGRGQLPOLP 1174
DB 2215 PLQANMPSTSRGRTMIHHPGVNNSSTSPVSKKGPPLKTPASKSPSEG---QTATTSF 2270
QY 1175 KOTLEERSAMDIEERNRQMLNKYKOVAGSDPRLQDYHSKYRSG---WDPHRGAD----- 1226
DB 2271 RGAKPSVKSSELSPVARQOT-----SQIGSS-----KAPSRSGSDSPSPRAQOPLSR 2318
QY 1227 TVSTKSSDSQSDVDAVASTSSASRFSSTSYMSVQSERGRGNKIKIVTFSSKMNQRMQVS 1286
DB 2319 PLOSPGRNISPGRNGISPPNKIJSOLPRTISSPSTASTKSSGSKMS--YTS--PGRQM--S 2373

```



```

Db 1066 ROSRMO-----TTYTESTDDKHLKFPHFQOE-----CVSPRSGANGSET 1112
QY 147 NCSYQOTFCARCGRGVSLRSKKNVMMVCLCKQOEILTRSGAMFNSSGNTLQOQDQK 206
Db 1113 N-----RVGNSHGINOVNSOLC-----QEDDYEDDKPTJNSERYSEEOHHE-- 1155
QY 207 PGLRNEBAPOEKKAHLKEPOFOGAPGLSVPAVEKGRAGHLTRDPTJKNCSGVKHOLA 266
Db 1156 -----EERPTVYSIKYNEKRVHVDOPIDYSL-----KYA 1184
QY 267 SDMPDRKRSBVSVDQNNRYEQSEEREDYSQVYPSDGTMPSPSDYADRRSQREFOYE 336
Db 1185 TDIPSOKOSFESFKSSGOSKTE-----HMSSENTSTPSSNAKRONOLHPSAQ 1237
QY 327 EPGHLYNDSNRGRHRSKE-----YIVD-----EDVE 355
Db 1238 SRSGOPQARATCKVSSINQETIOTYCEVEDTPICFSCSSLSLSADELICNOTTOEAD 1297
QY 356 SRDEXEROREEYQARYSDPNLARYPVKPOPYEOMRTHAEVSRARHRRSDVSLAN 415
Db 1298 SANTIQIAEIKETIGTRSNEDPVSEVPASQHPRTKSSRLQGS--SLSSSARKKAVEFSS 1356
QY 416 AELEDSTRISLMDRP-----SRQSVSEERRAMENQRSYSMERTREAO--- 459
Db 1357 GAKSPKSGAQTPKSPREHYOETPLMFSCRHSVS-----SLDSFESRSTASSVQSEPCSG 1412
QY 460 -----GGSYTPQRTSNISPPTRPRSPITLDRPDMRRADSLRKQHLDPSSAVRKTREK 513
Db 1413 MWSGIISPSDLPDPSGQTMPPSRKTPP-----PQTA--QTKRE- 1451
QY 514 METMLRNLSSDOSESVRPPPRHKSCKGKMROVSLSSSEEL-----ASTPEYTC 568
Db 1452 ---VPKNAPPAEKES-----GPKQAANVAQVOYVLPADOTLLHRLATESTPPGFSC 1502
QY 569 D-----DVEL-----ESESVESEKGSQ-KGKRT--SEOG 595
Db 1503 SSSLALSLDEPFIQDVELRIMPVQENDNGNETESPOKESNENQEKAEKXTIDSEKD 1562
QY 596 VLSDSN-----TRSEROKKRMYYGSHLEEDLEMESEPOIKDSGVTCSS 639
Db 1563 LLDSDDDDIIELECIISAMPTKSSRAKKPAQTASKLPPVARRPSDL-----PVYK 1616
QY 640 TTLNEHSHSDKHPYTWQPSKDGRL--IGRILN-----KRLKDGSVPRDSGAMLG 690
Db 1617 LLPSGNRLQOPKH-VSTFPGDMPRVYCEGTRPINFSTLSDLTIESPPLAAGEV 1675
QY 691 KVGCGKMTESGRLCATFKVKKGSLADTVGHLRPGDEVLENNGRLLQATFEVYNILE 750
Db 1676 R--GG--AOSG-----EFEKROTIPTEG--RSTDEA-----QGGKTSVTLPELD 1714
QY 751 SKPEQVELV-----SRPIGD-----IPRIPISTHAELESSSFEESOKMDRPSISVT 799
Db 1715 DNKAEGGILAEJINSAMPKSGSHKPRFVKKIMDOVO--QASASSSPANKNOLDGKKKPP 1773
QY 800 SPMSEGMULDVOPLSGOLSTIKLWFDKVGHOILVTLGAKDLPRSDGPRNRYVATYEL 859
Db 1774 SPVAP-----IPONTETRYV-----RKNAOSKNLNAERVF 1805
QY 860 PDRSDKNRRTKYVKLTLEPK-----WNQPTIYSPVH-----RR 893
Db 1806 SDNKSOKKONLKNNSKDFNDKLPNNEDRVRSGFADSPHHYPIETCTPYCFSSNDLSL 1865
QY 894 EFERMLEITLMDOKRVRESESEFGEIILETALLDDEPHWYKIQTHDVSSLPRLR-- 951
Db 1866 DFDDDDVLSR-EKALRKRAKKNKSEAKVTSHTELTLSQOASANKQA--IAQPIPNRQO 1922
QY 952 PSPYLPRLRQ-----HGESPTRLQR-----SKRISDESVDC----- 997
Db 1923 PKPILQKSTPQOSKODLPDKAATDEKIQNAIENTPVCFSHNSLSLSDIDQENNK 1982
QY 988 -----DGVGVSVDYRHHG-----RDLQSTLSVPEQVMSNHC 1040
Db 1983 ENPITEPTEPDSQEPSPKQASGFAPKSFHYEDTPVCSRNSSLSLSDISDLDLQEC 2042

```

```

QY 1021 SPGSGPHRVD---VIGTRMSPSAPPPQARNVEQCHRGTRATGHYNTISRMD-RHRVMD 1076
Db 2043 ISSAMPKKKRSRLKQNEKHP-----RNM-GGILLEDITLKLKDQRPSEGLSPD 2095
QY 1077 HYSDDRDRCDAAD-----RQ-----PYHRSRSTP 1101
Db 2096 SENFDKMAIQGANSIVSLQAAAAACLSQOASSDSISLSKGSISLGSFHLTPDOE 2155
QY 1102 QRPIL-----ERTTRRSRSERPDTN----- 1122
Db 2156 EKPTSMKGPRLKPGEKSTLETKEIESESGIKNGKVKYSLITGVNSSEISGOMQ 2215
QY 1123 -LMRSMPSLMGRSAPRPPALSRSHPRTGSV-----QTSSTSTPGGRRRQRLPOLP 1174
Db 2216 PLOANMPSISGRRTMTHIPGVNSSSTSPVSKKGPPLKTPAKSPSEB---QATATSP 2271
QY 1175 KGTLERSAMDIEERNRQMKLKKYKOVAGSDPRLDQDYHSKYRSG--WDPHRGAD----- 1226
Db 2272 RGAKPYSKSELSPVAKQT-----SOIGSS-----KAPSGSGRSDTPSKRPAQPLSR 2319
QY 1227 TVSTKSSDSDVSDVSAVSRSTSSARFSTSYNSVOSEPRGRNKISVFTSKMQRMGVS 1286
Db 2320 PLOSGRMSISPRGNISPPMKLSQLPRTSSPSTASTKSSGSGKMS-YTS--PGROM--S 2374
QY 1287 GKNTLKTSTISGDMCSLEKNGSOSDTAVGALGTSKGRRRSIGAKMAIVLSKRSRA 1346
Db 2375 QONLTKOTGLSKNASIIPRSBASAKGLNOMNNGANK-----VELSRMSTK 2423
QY 1347 SOLSTEGCGKKL-----RSTVORSTETGLAVEMNMNTROASRST----- 1388
Db 2424 SSGSESDRSERPVLYRQSTFIKEAPSPILRKLLESASFESLSRSPASPIRQAOITV 2483
QY 1389 -----DGSMNSYSS 1397
Db 2484 LSPSLPDMSLSTHSS 2498

```

```

RESULT 8
US-08-450-582-2
; Sequence 2, Application US/08450582
; Patent No. 6114124
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESSES:
ADDRESS: Banner & Wilcoff, Ltd.
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,582
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/452,655
; FILING DATE: 25-MAY-1995
; APPLICATION NUMBER: US 08/289,548
; FILING DATE: 12-AUG-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/741,940
; FILING DATE: 08-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.49964
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ. ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-450-582-2

```

```

Query Match      3.1%; Score 258.5; DB 3; Length 2843;
Best Local Similarity 17.9%; Pred. No. 2.7e-10;
Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69;

```

```

OY 30 SHLTERRKIIILAVMDROKKEEKESVLTKEEHNKAQPIQWFFSITELVNNVLPQQ 89
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1012 NMDDGDGLDLPINSLKYSDEQ---LMSGROSPQSNRMARPKHIIIE--DEIKSEQ 1065
OY 90 KQPNKEPQTKLHQEFEMKQEVKKMGEEQO---OQEQKGDAPTGICIKTKRACDGGH 146
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1066 RQSRNS---TTPVYTESTDDKHLKQPHFGQE-----CSPTRSKANSET 1112
OY 147 NCSYQCTKFCARCGHVSLSKSNKVMVCMCRKQOELLTKSGAMFYNGSGNTLQOPQKV 206
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1113 N-----RVGSHNGIINQVNSQISIC---QEDDYEDDKPTNYSERYSEEQHHE-- 1155
OY 207 PGRLNREAPQEKAKLHQPOFOGAPGDLSPAVAEKGRAGHGLTRDTLNGSGVKKQIA 266
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1156 -----EERPNTSYIKVNEKRIHQDPIDVSL-----KYA 1184
OY 267 SDMPDRKRKSPSVSRQDNRRYEQSEEREDYSOYVPSDGTMPRSPDYADRRSOREQFYE 326
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1185 TDIPSSQKQSFSEKSSSQSKTE-----HMSSSEMTSTPSSNAKQNLHPSSAQ 1237
OY 327 EPGHLNRYRDSNRGRHRSKE---YIVVD-----EDVE 355
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1238 SRSGQPOKAKATKVSSINQETIQYCEVEDPPICFSCSSLSSLSAEDEIGCNOTQGEAD 1297
OY 356 SRDEYERQREEREYQARIYSDPRLAITYPKAPQRYEQCMRIHAIVSRAHRRHSDVSLAN 415
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1298 SANTIQAIAIEIKRIGTSAEDPVSEVPAVSQHRTRKSSRLQGS-SLSSSESARHAAVEFS 1356
OY 416 ALEEDSRSLILMDRP-----SRQSVSEERRAMENQSYSMRTREAO----- 459
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1357 GAKSPSKSGAQRPKPEHYVQETPLMFSCRCTSVS---SLDSFESGISASSVQSECSG 1412
OY 460 -----GQSSYQORTSNHSPPTPRRSPPIPLDRPMRADSLRKQHLDPSSAAVAKTKREK 513
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1413 MVSGLISPDLDLPSPQOTPPSPRSKTPPP-----PQTA--QTKRE- 1451
OY 514 METMLRNDLSLSDGQSVAPPPRRPHKSKGGMKQVSLSSSEEL-----ASTPETTC 568
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1452 ---VFNKAPTAEKRDS-----GPKQAAVNAVAORQVLPDADTLIHFATESTPDQFSC 1502
OY 569 D-----DVEL-----SESVSEKGDQO-KGKRTK--SEOG 595
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1503 SSSLALSLDEPFIQDVLRIIMPVQENDNGNETSEQPKSENENQKAEKTIIDSEKD 1562
OY 596 VLISDN-----TRSERQKKRMVYGGHSLIEDLEWSEFQIKDQSGVDTCCSS 639
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 1563 LLDSDDDDIETLEECITISAMPTKSSRAKKPAQATASKLPPPVARKPSQL-----PYVK 1616
OY 640 TTLNEEHSKDHPIYTQDSKDGDL--IGRILLN-----KRLKDGVPKDSGAMGL 690
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1617 LLPQNRLOPOKH-VSLTPGDMPRVYCVGEGPINEFSTATSLDLTIESPPNELAAGEGV 1675
OY 691 KVGGMKTESGRLCAFTTKVKKGLADPVGHLPQDEVLWNGRLLQGTFFEEVNIILIE 750
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1676 R--GQ--AQSG-----EFEKRDITPTEG--KSTDEA-----QGGKISSVITPELD 1714
OY 751 SKPEQVELVY-----SRPID-----IPRIDSTHAQLESSSSSEFESQKMDRPSISVT 799
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1715 DNKADEGDIILACINSAMPKSKHPPRYKKIMQVQ-QASASSAPNNQDGGKKKKT 1773
OY 800 SPMSFGMLRDVPQFLSGQLSTKLMFDKVGHQLVITLLGAKDLPSREDGRPPRPYKITYL 859
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1774 SPVKP-----IPQNEYRTRV-----RKMDSKNNLNAERYE 1805
OY 860 PDRSDKNKRRTKVTKLEPK-----WNQFIYSPIVH-----RR 893
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1806 SDNKDSKQNLKNNSKDPNDKLPNNEDRYRGSFAFDSPPHHYPIGTPYCFGRNDSLSL 1865
OY 894 EFERMEITLMDQARVEESEEFLGEILIEETALLDEPHMYKLQTHDVSLELPR-- 951
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1866 DFDDDDVDLSR-EKAKELKAKENKSEAKVISHTELTSQOANNTQA--IAKQPINRQ 1922
OY 952 PPSYLPRLQ-----HGESPTRLQR-----SKRISDFVSQDCE-- 987
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1923 PKPILOKSTPQSSKIDIPDRQAATDEKLFQNTPEVCFSHNSLSLSIDIDENNRK 1982
OY 988 -----DGGVNVDYRNG-----RLOQSTLSPQVWSSNNC 1020
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1993 ENEPKTEFPDQSGEPKQPSAGYAPKSFHVEDIPVCFSRNSSLSSLSIDEDLLQEC 2042
OY 1021 SPQSPHARVD---VIGTRSMSPAPPROVYEQHGRGTRATGHYNTITSRMD-RHRVMD 1076
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2043 ISSAMPKKKKRPSRLKGDNEKHSF-----RNM-GGILGEDLLDKLDIQRPSHGLSPD 2095
OY 1077 HYSRDRDCCAAD-----RQ-----PYHRSRTE 1101
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2096 SENFDWKAIOGANSIVSLHQAAAACLRSQASSDSLSLKSGLSLGSPFHLPDQE 2155
OY 1102 QRPIL-----ERTTTRSRSSERPDTN----- 1122
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2156 EKPTTSNKGPRILRGEKSTLETKEIESEKIGKGGKVVYSLITGKYRSNSETIGQMKQ 2215
OY 1123 LMSMPSLMTGRSAPSPALSRSHPTGSV-----QTSPTSPTGGRGRQQLPOLPP 1174
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2216 PLQANMPISIRGRITIHLPGRNSSSTSPYSKKGPPLKTPAPKSPSG--QATATSP 2271
OY 1175 KGTLERSAMDLEERRQMKLKKYQOVAGSDPRLQDYHSTKTRSG--WDPHKGAD----- 1226
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2272 RGARPSYKSELSPVARQT-----SQIGGS-----KAPSSGSGRDSPTSPRAQOPLSR 2319
OY 1227 TVSTKSSDSDVDVSAVARTSSASRFSSTSYMSVQSEPRGRKTSVPTSKQUNQMGVS 1286
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2320 PIQSPGRKSTIPGNGISIPPKLSQLPRTSPSTASTSSSGSKGS-YTS-PGQOM--S 2374
OY 1287 GKULTKSTISIGMCSLEKNDGSDTAVALGTSIGKKRRSIGAKMVAIVGLSRKSSA 1346
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2375 QONLTQGTGLSKNASSIPRSESASKGLNMNNNGANK-----VELSRMSTK 2423
OY 1347 SOLQTEGGGKKL-----RSTVQSTETGLAVEKRNMTROASREST----- 1388
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2424 SSGSESDSESRPVLVROSTFIKEAPSPILRRKLEBSASFESLSPSRAPASPTRSQAQTPV 2483
OY 1389 -----DGSNNYS 1397
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2484 LSPFLPMKSLTHSS 2498

```

RESULT 9
US-08-450-582-7

```

Sequence 7, Application US/08450582
Patent No. 6114124
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THILVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Wilcoff, Ltd.
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,582
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/452,655
FILING DATE: 25-MAY-1995
APPLICATION NUMBER: US 08/289,548
FILING DATE: 12-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 08-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,49964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9299
TELEFAX: 202-508-9100
INFORMATION FOR SEO ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
US-08-450-582-7

```

```

Query Match 3.1%; Score 258.5; DB 3; Length 2843;
Best Local Similarity 17.9%; Pred. No. 2,7e-10;
Matches 304; Conservative 236; Mismatches 618; Indels 535; Gaps 69;

```

```

Db 1113 N-----RVGNHNGINQNVQSILC-----QEDDYEDDKPTNYSERYSEEQHEE-- 1155
QY 207 PRLRNEEAPQEEKAKLHEPOFOGABGDISVAVEGKRAHGLTRQDTIKNGSVKHOIA 266
Db 1156 -----EERTNYSIXYNEKRRHVDPIDYSL-----KYA 1184
QY 267 SDMPDRKRSRPSVSRQNNRYEEOSEEREDYQVYPSDGTMPRSPDYADRROREPOYE 326
Db 1185 TDIPISSOKOSFSKSSGSSKTE-----HMSSESTSTYSNAKKRONQLHPSAQ 1237
QY 327 EPGHLNRYDSNRGRHRSKE-----YIVDD-----EDVE 355
Db 1238 SRSGQPOKATCVSSINQTIQYCEVEDPICHRSCLSSLSASDEIGCQNTQOAD 1297
QY 356 SRPEYRQREERYQARYSDPNLARYPKVPEYEDMRHAENVSRARHERHSDVSLAN 415
Db 1298 SANTLQIAELKEKIGTRSAEDPVSEVPAVQHPPTKSSRLQGS-SLSESARHRAVERSS 1356
QY 416 AELEDRIILLRMDR-----SRORSYSEERRAAMENORSYEMERTREAO----- 459
Db 1357 GAKSPKSGAQTQPKSPPEKHYQETPLMFSCTSYS-----SLDSESSSIASSVSEPCSG 1412
QY 460 -----GQSSYPORTSNHSPPTRRRSEPIPLDRPDMRRADSLRKOHLLDPSSAVKRTREK 513
Db 1413 MVSGLTSPDLPSPQGTMPSPKSKTTPP-----PQTA--QTKRE- 1451
QY 514 METMLRNDLSLSQOSESVRPPPRPKSKGKMYOVSLSSEEL-----ASTPEYTC 568
Db 1452 ---VPRNKAPTAEKRSE-----GPKQAAVNAVQRYVLPDADTLHLFAETSPDSC 1502
QY 569 D-----DVEL-----ESESVEKGDQO-KGRKT--SPOG 595
Db 1503 SSSLALSLEDEPTIQDVELRIMPVQENDNGNETESQPKESNENQKEKKTITDSEKD 1562
QY 596 VLSDSN-----TRSEROKRMYYGGHSLBEDLEWSEPOIKDSQVTCSS 639
Db 1563 LDDSDDDDEIELECIISAMPTKSSKAKKPAQTAKLPPVARKPSQL-----PYVK 1616
QY 640 TTIINEHSHDKHPVWQPSKQDRL--IGRLILN-----KRLKGSVPRDSGAMIGL 690
Db 1617 LIPSQNRLOPKH-VSETPDDMDPRVYCVGCTPIEFSTATSLDLTIESPNEIAGBGV 1675
QY 691 KYVGMKTESGRICAFITKKGSLADTVGHLRPGDVLVLENGLLGATGEYYNITILE 750
Db 1676 R--GG--AOSG-----EKKRDTIPTEG--RSTDA-----OGKTSSTIPELD 1714
QY 751 SKPEQVELVY-----SRPGD-----IPRIDSTAOLESSSSSESQMDPISVT 799
Db 1715 DNKAEGDILAEICINSAMPKSGHKKPRVAKKINDQV-QASASSANKNQDLGKKKPT 1773
QY 800 SPNSPGMLRDVPOFLSGQLSIKLMFDKVGHQILVTLGANDLPSEDEGPRPNPYVKIYFL 859
Db 1774 SPVKP-----IPONTEYRTRV-----RKNADSKNNLNAERVF 1805
QY 860 PDRSDKNKRRTKVTKTLEK-----WNGFTIYSPVH-----RR 893
Db 1806 SDMKDSKKOMIKNNNSKDFDNKLPNNEDEVGRGSAFDPSPHHTPIEGTPYCFSRNDSLSSL 1865
QY 894 EFERMLELITLMDQARRESEFLGILLETALDLDEDPHWKLTQHDVSSILPLR- 951
Db 1866 DFDLDDVDLSR-EKAEIRKAKENKESAKYTSIHLETSNOOSAKTKA--IAKPIINRGQ 1922
QY 952 PSEYLPRLQ-----HGESPTRRIQ-----SKRISDEVSVDYCE--- 987
Db 1923 PKPTLQKSTFFPOSSKDIPDRCATDEKLQNFALIENTPVCFSHNSLSLSDIDQENNNK 1982
QY 988 -----DGVYGVSDVRHNG-----FDLQSSVLSVDEQVMSSNNHC 1020
Db 1983 ENEPITEPEPDSQGESEKQOASGVAPKSFHVEDTPVCFSRNSSLISLSDIEDLLQEC 2042
QY 1021 SPFGSPHRYVD--VIGRTRSMSPAPPPQNRNVGHHGTATGTYNTRISMD--HRYVMD 1076
Db 2043 ISSAMPKKKKPRSLKGDNEKHP-----RNM-GGILGEDTLTLDKTIQRPDSHGSLSPD 2095

```

```

QY 1077 HXSRRDRCCEAD-----RQ-----PYHRSKTE 1101
Db 2096 SNFPMKALQOEGANSTVSSIHQAAAAACLSRQSSDSSTLSLSKSGISLSPFLITDQE 2155
QY 1102 QRPRL-----ERTTSSRSSERPDTN-----1122
Db 2156 EKPFISNKGPRILKPGCEKSTLFTKRIEESKGIKGGKVKYLITGKVRNSSETISGOMQ 2215
QY 1123 -LMBRSPSLMTGRSAPSPALSLSHRTGSV-----QISPSSTPGTGKRGRLPQLP 1174
Db 2216 PLQANNPSTISRGTMHIGVNRSSSTSPVSKKGPPLKTPAKSPSEG---QTATPSP 2271
QY 1175 KGLTERSAADIERNRHOMKLNKYQVAGSDPRLQDYHSKYRSG---WDPHRCAD-----1226
Db 2272 RBAKPSVKSELSPVARTQ-----SQIGSS-----KAPSRGSRNSTSTSRAPQQLSR 2319
QY 1227 TVSTSSSDSDVDSAVSTSSASRFSSTYSMSVOSERPGRNKRISVFTSKMONRQMGVS 1286
Db 2320 PLOSFGNRNISFGNRGISPNNKLSQLPRTSSPSTASTKSSSGSKMS-YTS--PGRQM--S 2374
QY 1287 GNLTKSTISIDMCLEKNDGSDPTAVCALGTSGKKRRSSIGAKKVAIVGLSKRSRA 1346
Db 2375 QONLTKOTLSKNASSIPSESASKGLMNNNGANKR-----VELSRMSTK 2423
QY 1347 SOLSOTEGGKKL-----RSTVQSTETGGLAVERMMMTROASREST-----1388
Db 2424 SSGSESDBREPVLVQSTFIKAPSPILRKLEESASFESLSPSSKPAAPTSSQAQTPV 2483
QY 1389 -----DGSMSYSS 1397
Db 2484 LSPSLPDMSLSTHSS 2498

```

```

RESULT 10
US-08-821-355A-7
; Sequence 7, Application US/08821355A
; Patent No. 5851775
; GENERAL INFORMATION:
; APPLICANT: Barker, Nick
; APPLICANT: Clevers, Hans
; APPLICANT: Korinek, Vladimir
; APPLICANT: Morin, Patrice
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; APPLICANT: Sparks, Andrew
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
; TITLE OF INVENTIONS: Interact to Prevent Cancer
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,355A
; FILING DATE: 20-MAR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32,145
; REFERENCE/DOCKET NUMBER: 1107.05064
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 97430 BMB UT
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2973 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO. 5851775e
; US-08-821-355A-7

```

```

Query Match 3.1%, Score 258.5; DB 2; Length 2973;
Best Local Similarity 17.9%; Pred. No. 2.9e-10;
Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69;

```

```

QY 30 SHLFEERKILAVMDROKKEEKESVLIKIEHKAPOTQMPFSGITELVNNVLOPOQ 89
Db 1012 NHMDNDDELDTPIVNSLYKSDQE---LNSGRQSPSQNERWARKHTE--DEIKQSQ 1065
QY 90 KQPNKEPOTLIHQEFEMVKKQVKKMGESQO---QOEKGDAPPCGICHTKPKADGCH 146
Db 1066 RQSRNQ-----TTPVYTESTDDKHLKFPNFGQOE-----CVSPYRSRGANGSET 1112
QY 147 NCSYCOFTKFCACGCGRVLRSNKVMWVCNLCKRQOEILTKSGAMFYNSGNTLQOPDOKV 206
Db 1113 N-----RVSGNHGINQVNSQSLC-----QDDYEDDKPTVYSEYSEEDHE-- 1155
QY 207 PRGLRNEEAPQEKAKLHEDQFOGAPGDLVPAVEKGRANGLTRQDTIKNGSVKHOIA 266
Db 1156 -----EERPTNYSIKYNEEKRHYDQPIDYSL-----KTA 1184
QY 267 SDMPDRKRSPSVSDQNRKRYEQSEFEDYQVVPDGTMPSPSDYADRKQREPOFYE 326
Db 1185 TDIPSSQKQSFSSKSSGQSGKTE-----HMSSESTNTSPSSNAKRONLHPSSAQ 1237
QY 327 EPGHLNVRDSNRGRHNSKE---YIVD-----EDVE 355
Db 1238 SRSGQPOKATCKVSSINQETIQTICVEDTPICFRCSSLSLSAEDIGCNOTTOEAD 1297
QY 356 SRDEYERQREERYQARYRSDPNLARVPVKQPYEQMRINAEVSRARRHRSDVSLAN 415
Db 1298 SANLQLINEIKETIGTRAEADVSEVPVAVSQHPRTKSSRLQGS--SLSSSAHKAVERSS 1356
QY 416 AELEDRIISLRMDRP-----SRORSVERRAAMENQSYMERIREMO-----459
Db 1357 GAKSPSKGAQTPKSPREHYVQETPLMFSRCTSVS---SLDSFESRSTASSVQSEPCSG 1412
QY 460 -----GQSSYPORTSNPSPTPRRSPIPLDPRMRADSLAKQHHLDPSSAVRKTREK 513
Db 1413 MVSGLISFDLPDSEPGQVMPSPSRKTPPP-----PPQTA--QTKRE- 1451
QY 514 METMLRNDLSLSDQSESVRPPPRHKSCKGKMROVLSLSSSEEL-----ASTPEYTSK 568
Db 1452 ---VPKNKAPAEKRES-----GPKQAAVNAVAQVQVLPDADVLHFRATSPDGFSC 1502
QY 569 D-----DYEL-----ESESYSKQGSQ--KGAKKT--SEOG 595
Db 1503 SSSLATSLDEPFIQKDELIRIMPVQENDNGENTESQPKESNNOEAEKRTIDSEKD 1562
QY 596 VLDSN-----TRSEQRKRMVYVGHSLIEDLEMSPOIKNSGVDTCSS 639
Db 1563 LDDSDDDDIILEECIISAMPTSSRAKKAQTAASKLPPVAKKPSOL-----PYK 1616
QY 640 TTLNEHSHDKHPVYTWOPSKRGDRL--IGRIILN-----KRLKDSVPRDSGAMGL 690
Db 1617 LLPQONRLQPKH--VSTFRGDDMPRVYVEEGRPIFNSTATLSLDTLTPSPRELAGEVY 1675
QY 691 KVGCKMTESGRICAFITKVKKGLADTVGHLRPGDEVLEWNGRLLOGATFEVYNIILE 750
Db 1676 R--GG--AQSG-----EFEKRTIIPTEG--RSTDEA-----QGGKTSVITPELD 1714

```

```

QY 751 SKPEQVELV-----SRPIGD-----IPRIPDTHAOLESSSSSESQKMDRPSISVT 799
D 1715 DNKAEEGIIIAECIAMSAPKSGSHKPRKKTIMDOVO-QASASSAPNKNOLDGKKKFT 1773
QY 800 SPMSPGMLRDVPOFLSGQLSIKLMWDKVGHOILVITLAKADLPREDGPRPRPVKVIYFL 859
D 1774 SPVKP-----IPONTYERTRV-----RKNADSKNNLNAERVE 1805
QY 860 PDRSDKNKRRTKYKTLERK-----WNQFTIYSPVH-----RR 893
D 1806 SDNKSQKRLKNKNKNDKLPNNEDRVGSAFADSPHHYPRIECTPYCFGRNDLSL 1865
QY 894 EFERMELEITLMDQARVREESEFLEIIELEFALDDEPHMYKLOTADVSSLPLPR-- 951
D 1866 DFDDDDVULSR-EKAELEKAKENKSEAKVYSHTELTSTQOSANKTQA--IAKQPIRNGQ 1922
QY 952 PPSYLLPRRQL-----HGESPTRLQR-----SKRISDSVSDVCE--- 987
D 1923 PKPILQKOSTFPOSKDIIPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENNNK 1982
QY 988 -----DGAVVSDYRNG-----RDLQSTLSVPEQVMSNNC 1020
D 1983 ENERIKETEPPOSGEPKPOASGYAPKSFHVEDPVCFSRNSLSLSISDEDDLQRC 2042
QY 1021 SPSPGPHRVD--VIGRTSRMSPSAPPPQANVEQHGRTGRATGHYNTISRMD-RHRVMD 1076
D 2043 ISSAMPKAKKKSRLKGDNEKHSPP-----RNM-GGILGEDLILDKIDRPPDSEHSLSD 2095
QY 1077 HYSDDRDRDCEAD-----RQ-----PYHRSRSTE 1101
D 2096 SENFWMKAIQEGANSIVSLHQAAAAACLSRQASSDSLSLSLKGSLISGPHLPPDQE 2155
QY 1102 QRPPL-----ERTTIRSSSERPDN----- 1122
D 2156 EKPFTSNKGPRLKPKERKSTLETKKIESEKGIKGGKVVYKSLITGKVSNSEISQCMQ 2215
QY 1123 -LMRSPMLTGRSAPSPALSRSHPRTGSV-----OTSPTPTGTRGRQLQLDP 1174
D 2216 PLOAMPSPISRKRTIHIIPGVANSSSSTSIPVSKGPPPLTPAKSPSECG---QTATTSB 2271
QY 1175 KGLTERSAMDIERNRQMKLNKYQOVAGSDPRLDQYHSGYRSG--WDPHRCAD----- 1226
D 2272 RGAPSVKSELSVPAQRQ-----SQIGSS-----KAPSRSGSDRTPSRRAQPLSR 2319
QY 1227 TVSTKSSSDVSDVAVRTSASRPSSTSYSVOSERPGRNKISVFTSKMONRQMGVS 1286
D 2320 PLOSGRNSISPGRNGISPPNKLSQLPRTSSPSTASTKSSGSGKMS-YTS--PGRQM--S 2374
QY 1287 GKNLKRTSISGDMCSLEKNDGOSQDTAVGALGTSGKKRRRSIGAMVAIVGLSRKRSRA 1346
D 2375 QONLTKQTLGSLKNASSIPRSSESAGLONMNGCANKR-----VELSRMSSTK 2423
QY 1347 SOLSOTEGGKKL-----RSTVORSTETGLAVEMRMNMTROASREST----- 1388
D 2424 SSGSSDSRSEVRVLKROSTFIKEAPSPILRRKLEBASPFESLSPSRPASPTPSQAQTPV 2483
QY 1389 -----DGSMSYSS 1397
D 2484 LSPSLPDMSLSTHSS 2498

```

RESULT 11

US-09-003-687A-7
Sequence 7, Application US/09003687A

Patent No. 5998600

GENERAL INFORMATION:

APPLICANT: Barker, Nick

APPLICANT: Clevers, Hans

APPLICANT: Korinek, Vladimir

APPLICANT: Morin, Patrice

APPLICANT: Kinzler, Kenneth

APPLICANT: Vogelstein, Bert

APPLICANT: Sparks, Andrew

```

? TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
? TITLE OF INVENTION: Interact to Prevent Cancer
? NUMBER OF SEQUENCES: 11
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Banner & Witcoff, Ltd.
? STREET: 1001 G Street, N.W.
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20001
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/003,687A
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/821,355
? FILING DATE: 20-MAR-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Kagan, Sarah A
? REGISTRATION NUMBER: 32,145
? REFERENCE/DOCKET NUMBER: 1107.05064
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-508-9100
? TELEFAX: 202-508-9299
? TELEX: 97430 BMB UR
? INFORMATION FOR SEO ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2973 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: No. 5998600e
US-09-003-687A-7

```

```

Query Match 3.1%; Score 258.5; DB 2; Length 2973;
Best Local Similarity 17.9%; Pred. No. 2,9e-10;
Matches 304; Conservative 239; Mismatches 618; Indels 535; Gaps 69;

QY 30 SHLTREERKIIILAVMDRQKKEERQSVLYKTEHKAQPTQMPFSGITELVNNVLPQOQ 89
D 1012 NHMDNDGELDTPINYSILKYSDEQ---LNSGRQSPSONERWARPKHIE--DEIKQSEQ 1065
QY 90 KQPERKEPQTRLHQQEFMYKQYVKKMGESEQ---QOEQKGDAPTCGICHTKTFADGCGH 146
D 1066 ROSRNQS-----TTYVYVTESTDCKHLKFPHFQOQ-----CVSPYRSGANGSET 1112
QY 147 NCSYCQTKFCARCGGRVSLRSNKVMVYCNLCRQOEILTKSGAMEYNSGNTLQOPDOKV 206
D 1113 N-----RVGSNNGINQNVQSGLC---QEDVEDDKPTIYNSIRYSPEEQHEE-- 1155
QY 207 PRLGLNEAPQOEKAKLHEDQFOGAPEDLSVPAVEKGRANGLTRQDTIKSGYKHQIA 266
D 1156 -----BERPTNYSIKYNEEKRHVDQPIDYSL-----KYA 1184
QY 267 SDMPDRKRSPSVARDQNRREYQOSEEREDYSQYVPSDGTMPRSPDYADRRSOREPQYE 326
D 1185 TDIPSSQKOSTFSTKSSGQSSKTE-----HMSSESTSTPSNAKRONQLHPSAQ 1237
QY 327 EPGHLNRYDSNRGRHRSKE-----YIVDD-----EDVE 355
D 1238 SRSGQPKAATCKVYSINQETIQTYCVEDPTICFSRCSLSLSAEDIGCNOQTQAD 1297
QY 356 SRDYEERQREBEYQARYRSPNLARYVYKQPYEQRITAEVYSRAHRRHSVSLAN 415
D 1298 SANTLQIAETKEKIGTBSAEDPVSEVPAVSQHPRTKSSRLQGS--SLSESAKRAHVEPSS 1356
QY 416 AELEDSRISILRMKRP-----SPQRSVSERAAEMNQRYSMEETREAO----- 459

```

Db	1357	GAKSPSCGAQTKPSPENHYQETPLMFIRSTVS-----SILDSFENSIASVSGEPCSG	1412
QY	460	-----GQSSYPORTSNHSPPTPRASRSLPLDRPDMRADSLRKOHILDPSSAVKTKREK	513
Db	1413	MVSGIISPOLPSPSPGQTMPPSRKSTPP-----PPQTA--QTKRE-	1451
QY	514	MEMLRNDLSLSDQSESVRPPPPPHKSKGKKRGVSLSSSEEL-----ASTPEYTCG	568
Db	1452	---VPKNKAPTAKRES-----GPKAAVNAVQROVLEPDAJTLHFAETSPDDEGSC	1502
QY	569	D-----DVEL-----ESESYSSEKGDQO--KGKTKT--SEOG	595
Db	1503	SSSLASLSDPEFIQKQVDELIRMPVOENDNGETSEEOQESKENENQEKAEKTIIDSEKD	1562
QY	596	VLSDSN-----TRSEORCKRMYYGGHSLDELEMSERPOLKDSVDTCSS	639
Db	1563	LDDSDDDDEILEEFCILSMPTKSKKAKKAPQATASKLPPVPARKPSQL-----PVYK	1616
QY	640	TLTNEHSHDKHPVTWQPSKDDRL--IGRIILN-----KRLKGSVPROSGAMIGL	690
Db	1617	LTPSQNRLOPKH-VSEFTPGDDDMRYVCVEGTFINFTATSLSDLTIESPPNELAABGV	1675
QY	691	KVYQKMTSGRCAFITTKKKKSLADTVHRLRGDEVLEMGRLLOGAFEEVYNIILE	750
Db	1676	R--GG--AOSG-----EFKKDTIPTEG--RSTBA-----OGKTSVITIELD	1714
QY	751	SKPEPOVELV-----SRPGD-----IRIPDSTHAOLESSSSPESQKMDPSISVT	799
Db	1715	DKKAEBGDLAECINSAMPGRGKHKFRVYKKINDQVO--QASASSAANKKQLODKKKKPT	1773
QY	800	SPKSPMLADVPOFLSGQLSTIKMFDKVGHQILVTLIGKADPLESREDGRPNRYVYKLYL	859
Db	1774	SPVKP-----IPONTEYRTRV-----RKMAKSNMLNAERVE	1805
QY	860	PPRSDKNRRRTVTKLTLPK-----WNOTFISPVH-----RR	893
Db	1806	SNNKSKKONLKNNSKDNDFDKLPNNNDKRVGSAFSPHHYITIEGTPYCFSRNDSISL	1865
QY	894	EFERMELETLMDQAVRESESEFELIILEETALLDEPHYKLIQTHDVSSPLPR--	951
Db	1866	DDDDDDLSLR--EKALIRAKEREKESEAKVTSHTELTSNQSANKTQA--IAPQINRGQ	1922
QY	952	PSPYLPRRL-----HGRSPRIRLQR-----SKRIDSEKDYDCE---	987
Db	1923	KPPILOKQSTPFOOSKDIPDRGATDEKLQNEAIENTPVCFSHNNSISLSDIDQENNNK	1982
QY	988	-----DGVGVVSDYRHNG-----RDLOASTLAVPEQVMASNHG	1020
Db	1983	ENEPKIKETEPPOSGEPRKQASGVAFKSFHVEDTPVCFSRNSSLSSLSDSEDLLOEC	2042
QY	1021	SPSGSPHARD--VIGRTSMSPSAPPPORANBOGHRGTRATGHTVITSM-D--RHRVMD	1076
Db	2043	ISSAMPKKKPSRLKGDNEKHP-----RNM--GGLIGEDLTLDKDIORPDSEHGLSP	2095
QY	1077	HYSSDRDRDCEAD-----RO-----PYHRSRME	1101
Db	2096	SENFPMKMAIQEBANSTVSLHAAAAACLSROASSDSLSLSKSTISLGSPHILTPDOE	2155
QY	1102	QRPLL-----ERTTTRSSERPDTN-----	1122
Db	2156	EKPFTSNKGPRILKGEKSTLETKKIESSEKIGKGRKYKSLITGVANSNSELISQMKO	2215
QY	1123	LMRSPMLMGRSAPPSALSRHPRGSV-----QTSSTPTGTRGRGOLPOLPP	1174
Db	2216	PLQAMPPTISRKRTIHLIPGVNSSSSITSVPVKKGPILKTPAKSPSEG---QTAATSP	2271
QY	1175	KGTLERSAMDIEERNRQMKLNYKKOVAGSDPRLQEDYHKKYASG---WDPHRCAD---	1226
Db	2272	RGAKPSVSKSLSPVARQT-----SOLIGSS-----KAPSGSGNDSTPPSKPAQOPLSR	2319
QY	1227	TVSTKSSSDVSDVSAVSKTSSASRSTSTSVSVOSERPGRNKRKISVFTSKMONRQMGVS	1286

```

Db 2320 PIGSPGRNRISPGNGNGISPPNNKLSQLEPRITSSPETASTKSSGGGKMS-YTS--PGHM--S 2374d
QY 1297 GKNLTKSTSIISGDMCSLEKNDGSDSPDPAVALCTSGKKRRSSIGAKWAIVGLSRKSRSA 1346f
Db 2375 QONLTQGTGLSKNASSIPRSESASGKLNQNNNGNCANKR-----VELSMNSTK 2423h
QY 1347 SOLSOTEGGGKKL-----RSTVORSTETGLAVEMNMNTROASRST----- 1388i
Db 2424 SSGSESDSRSEPPVLYRQSTFIKEAPSPITLRKLEESASRSESPSSRPASPTRSQAQTPY 2483j
QY 1389 -----DGSNMYSYS 1397l
Db 2484 LSPSLPDMSLRSTHSS 2498m

```

```

RESULT 12
US-09-136-605-7
; Sequence 7, Application US/09136605A
; Patent No. 6140052
; GENERAL INFORMATION:
; APPLICANT: He, Tong-Chuan
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert.
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC Interact to
; TITLE OF INVENTION: Prevent Cancer
; FILE REFERENCE: 1107.75741
; CURRENT APPLICATION NUMBER: US/09/136,605A
; CURRENT FILING DATE: 1998-08-20
; EARLIER APPLICATION NUMBER: 08/821,355
; EARLIER FILING DATE: 1997-03-20
; EARLIER APPLICATION NUMBER: 09/003,687
; EARLIER FILING DATE: 1998-01-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2973
; TYPE: PRRT
; ORGANISM: Homo sapiens
; US-09-136-605-7

```

	Query Match	3.1%: Score 258.5; DB: 4; Length 2973;
	Best Local Similarity 17.9%: Pred. No. 2.9e-10;	
	Matches 304; Conservative 238; Mismatches 618; Indels 535; Gaps 69;	
QY	30 SHLTTEERKILLAAVDRCKKEEKQSVLKTKEHKAQPTQWFEPFGITTELVNNALQEQ	89
Db	1012 NHMDNDGELTPTINTSLKYSDEQ-----LNSGRQSPSQNEHARPKHILE--DEIKQSDQ	1065
QY	90 KQPNKEDEQRTLHOQEFMYKQYQVKMGESEQ--QQQKQKDAPTCGICHTKTFADGGCH	146
Db	1066 KQSNQNS-----TTTPYLTSTDPDKHLKFPQPHQGOE-----CYSPIYSRKNGSSET	1112
QY	147 NCASYQCTKFCARCGRVSLRSNKVWVYCNLCRKQOELLTKSGAFYNGSGWTLQOPDKV	206
Db	1113 N-----RVGSSNGINGQNSQSIC-----QEDVDYEDKCPNNYSRYSYEEQRHEE--	1155
QY	207 PGLRLNEEAQEQKAKLHEQPFQOGAPGDDLSVPAVERGGAHGLTRQDTIKNGSGVHOIA	266
Db	1156 -----EERPLNYSIKYNEERKHVDQPIDSL-----KYA	1184
QY	267 SDMSDDKRRSPSVSRDQNNRRYEOSEEREDYSQYVPDQTMRSPDYADRRSOREPQYQE	326
Db	1185 TDIPISSQKQSFSPFKSSSSGQSKTE-----HMSSESENTSTPSSNAKRONQLHPSSAQ	1237
QY	327 EPGHLNTRDSNRRGRHRSKE-----YIVDD-----EDVE	355
Db	1238 SRSQGPQKATCKKSSINQETIQTYCEVDPIPCPSRCSLSLSSADETGQNTQDAD	1297
QY	356 SRDPEHQREEREQYQARYSDPNLARFPVQRPQRYEEDQMRTIAEYSRAHRRHSDVSLAN	415
Db	1298 SANLQJALETKEKIGTNSADPVESEVPAYVSOHPTKSRSLQGS--SLSESARNAHAYESS	1356

QY	416	AELEDSISILLRMDRP-----	SRQSVERRAMNQSVMERTREAO-----	455	
Db	1357	GANSFSGAQPKSPREHVOETPLMESCTSVS---	SLDSFERSIASSVOSPCSG	1412	
QY	460	-----GQSSYPORTSNHSPPTRRSPILDRPDMRADSLKQHNHLPPSAVKTREK		513	
Db	1413	MVSGIISPDLPDPSQOTMIPMSKIPPP-----	PPQTA-QTKRE	1451	
QY	514	MEWLNRDLSIDQSSQSVRRPPRPHKSGKGMQVSLSSSEEL----	ASTPEYTC	568	
Db	1452	-----VPKKAKPTAKRES-----	GPKQAAVNAVAORVQVLPADLTLLHFATIESTPDGSC	1502	
QY	569	D-----DVEL-----	ESEVSQSGSO-KGKRKT--SPQG	595	
Db	1503	SSSLSLSLDPEPTQKQVELRIMPPVQENDNGNETESQPKSNNOKEAREKTIIDSKD		1562	
QY	596	VLSDSN-----	TRSEROKKRMYYGSHSLBEDLEMSFOIKDQGVDTSS	639	
Db	1563	LDDSDDDDDLEIEELCLISAMPTKSRKAKKAQAASKLPPVAKKBPOL-----	PYVK	1616	
QY	640	TTINEESHSDKHPTWOPSKQDRL-IGRILLN-----	KRLKDSVPRDGCAMGL	690	
Db	1617	LPLPSONLQPKH-VSEFTPGDDMPRVYQVEGPIINFSTATSLDLTIESPMLAAGCV	1675	1755	
QY	691	KVNGCKTEGSRCLAFITKYKKSGLADYGHLRPDQDEVLEWNGRLQGATEEYVNIILE	750		
Db	1676	R--GG--AQSG-----	BEKKDITPTG--RSTDEA-----	QGGKTSVITPELD	1714
QY	751	SKPEQVELV-----	SRPIGD-----	IPRIPDTHAOLESSSSPESQKMDRISVT	799
Db	1715	DNKAEBDILAEICNSAMKRGKSHKFRKTKIMDOYQ-QASASSAPKKNQDQKKKKPT		1773	
QY	800	SPMSGMLRDVPOFLSGOLSIKIMFDKVGHOLIVTILGAKDLPSREDGRPRNYYKIYEL		859	
Db	1774	SPVKP-----	IPONTEYRTRV-----	RKNDSKNNINAEVVF	1805
QY	860	PDSMDKRRKTKYVKTLEPK-----	WNQFTISPVH-----	RR	893
Db	1806	SDNKDSKKQMLKNSKDFNDKLPNNEDRYRGSPAFDPSPHNYITGTPYCESRDNLSL		1865	
QY	894	EFERMELEITLMOQARVRESESEFLIELIETLALDDEPMWYKQHNDSVLPFR-		951	
Db	1866	DFDDDDVDLSR-EKAEILKAKKEKESEAKVTSHTELTSMQSANKTQA--	IADQFINQO	1922	
QY	952	PSPYLPRQL-----	HGSEPPRRLQR-----	SKRISDEVSYDDE--	987
Db	1923	KPKILOQSFPPQSKNDIPRGATDEKIQNAIENTPYCFSHNSLSLSDIOENNNK		1982	
QY	988	-----DCGVVSDYRANG-----	RDQSGTSLSVBQVWSSMHC	1020	
Db	1983	ENEFIKETEPDQSGERSKQASGVAPKSFHEDPVCFSRNSLSISEDDLQEC		2042	
QY	1021	SPGSPSHRV-----	VIGETRSWSPSAPPQBNVBOGHRGRATGHNITSRMD-RHRVMD	1076	
Db	2043	ISSAMPKPKPRSLKQDNEKHSF-----	RNM-GGILGEDLTLDIKDQRPSEGLSPD	2095	
QY	1077	HYSSDRDRDCEAD-----	RO-----	PYHRSSTE	1101
Db	2096	SENEDKAIIOGANSIYVSLHQAAAACLSROASSDSILSKGISLGSFHLTPROE		2155	
QY	1102	QRPLL-----	ERTTSSRSSERPDTN-----		1122
Db	2156	EKPPTSXKSGPRILKPGKSEKTLTKIESKGIKGGKVKYSLITGKVSNGSEIGOMQO		2215	
QY	1123	LMSMSLMTGSGAPPSPLSLSRHRTQSV-----	OTSPSSPPTGGRGRQRLQPLR	1174	
Db	2216	PLQANMPSISRGRTMHIPEVRNSSSTSPVSKGKGPPLKTPASKPSSEG----	QIATTPSP	2271	
QY	1175	KGTLEBSAMDIEERNQMLNKYKQVAGSDPRLQEDYHYSKYSG----	WDPHNGAD-----	1266	
Db	2272	RGAPSVKSELSPVART-----	SQIGGSS-----	KAPSGSGRDSIPSPKPAQPLSR	2319
QY	1227	TVSTKSSSDVDVSAVSTTSASRFSSTSYMSVOSERPGRNKRITVFTSKMQRQMGVS		1286	

FILING DATE: 12-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
TELEPHONE/DOCKET INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-289-548A-2

Query Match 3.0%; Score 250.5; DB 1; Length 2843;
Best Local Similarity 18.0%; Pred. No. 1,1e-09;
Matches 305; Conservative 235; Mismatches 620; Indels 535; Gaps 70;

30 SHLTBEERKILLAVMDROKKEEKEOSVLTKEEHRKQPTQWPFSGITELVNVNLOPQ 89
1012 NHMDNDGELDPTINYSKXSDQ---LNSGRQSPQSNRMAPKHIIE--DEIKQSEQ 1065
90 KONEKEPOTKLHQEFMYKEQYKKMGESQO---OOEGKDAFTGICHKTRFADCGH 146
1066 KOSRNO---TTPYVTESTDDKHLKQPHGOE-----CVSPKSRKANSSET 1112
147 MCSYCOTKFCARCGRVSLSNRKVMVNCILRKQOELLTKSGAMFVSGSNLTLOPQKV 206
1113 N-----RVGSHNGINQNVQSGLC-----QEDDYEDDKPTINYSERSEEQHEE- 1155
207 PGLRNEAPQOEKAKLHEOPQOGAPGDLSPVAVEKGRAGHLTRQDTINKSGVKKQIOA 266
1156 ---EERTNYSIKYNEKRRHYDQPIDYSL-----KYA 1184
267 SDMPDRKRSPSVSRQNRREYEOSEEREDYSQYVPSDGTMPRSPDOVADRROREPOFE 326
1185 TDIPSSQKQSFSSKSSGSSKTE-----HMSSSESTSTPSSNAKKRONLHPSSAQ 1237
327 EPCHLYRDSNRGRHRSK---YIVDDDV-----ESRDEYERQREERYQ 370
1238 SRSGQPKAKATCVSSINQETIOTYCEVEDPFCRCSLSLSLSADEIGCNLTQDEAD 1297
371 A-----RYRSDPNLARIYKPYQYEEQMRHAEYRARHERHSDVSLAN 415
1298 SANTLOIAELIKGIGTRSAEDPVSEVPVAVSOHPRTKSSRLQGS--SLSESARHRAVEFP 1356
416 AELEDSRISLLRMDRP-----SPORSYSERRAMENORSYSMERTREAO----- 459
1357 GAKSPKSGAQOTKSPENHYQETPLMFSTKTSVS---SLDSFESSIASVSEPCSG 1412
460 -----GQSSYPORTSNHSPTEPRSRSPILDRPDMRADSLRKOHNDPSSAVKTKRK 513
1413 MVSGLISPSOLPSPQGTMPRSKTPP-----PQTA--QTKRE- 1451
514 METMLRNDLSLSQOSESVPRPAPRPHKSKGKMKROYSLSSSEEL---ASPEPTSC 568
1452 ---VPRNKAPTAEKRRES-----GPKQAAVNAVQROVYLPDADTLHLFATESPPDGSC 1502
569 D-----DVEL-----ESESVESEKDSQ--KGRKRT--SPQ 595
1503 SSSLALSILDEPTQKQVELRIMPVOENDNGNETSEQKESNENOEKEKTIIDSKD 1552
596 VLSDSN-----TRSEROKKRMYYGGHSLDELDLWSPDLKSDGVTCSS 639
1563 LLDSDDDDEILEECITISAMPTKSSRKKGKPAQTASKLPPVARKPSQL-----PYK 1616
640 TITNEHSHSDKRPVYTWQSPKQDRL--IGRILL-----KRLKQGSVPDROGAMIGL 690
1617 LIPSQNLQPKH--VSFTPDMDMFRVYCVEGTPIINFSTATSLSDLTIESPENELAGGV 1675

691 KVGCGKMTESGRICAFITKVKKSLADYVGHLPQDEVLEWNGRLQCATFEERYNILE 750
1676 R--GG--AQSG-----EPEKRDITPTG--RSTDEA-----QGKTSSTVITPELD 1714
751 SKPEQVELVY-----SRPIGD-----IPRIPDSTHAQLESSSSSEFQKMDRPSISVT 799
1715 DNKAEEGDILAECLINSAMPKGSKHPFVKKIMDQVO--QASASSAPNNQDLGKKRKP 1773
800 SPMSPGMLRDYQPLSGQLSTKLMFQDKVGHOLIYVILGAKDLPSNEDGPRRPYKITYL 859
1774 SPVKP-----IPQNTYRTRV-----RKNADSKNNLNAERYE 1805
860 PDRSDKNKRRTKTKVTKLEPK-----WNQFTYSPVH-----RR 893
1806 SDNDSKRONLKNNSKDNDRKLPNNEDRVGSAFADSPHHTPIEGTPTPCFSRNDSSL 1865
894 EFERMLEITLMDQARVBESEFICETILIEETALLDDEPHMYKIQTHDVSLPLPR-- 951
1866 DEDDDDVLSR-EKAEILRAKENKSEAKVTSHTELTSSNOQANKTQA--IAKQPINRQ 1922
952 PSPIYLRQL-----HGESPTRLQR-----SKRISDEVSIDCE----- 987
1923 PKPILQKSTFPQSKDIPDRGAATDEKLQNPALIENTPVCFSHNSLSLSDIDQENNK 1982
988 -----DGVGVSDYRHNG-----RLQSSSTLSVPPQVMSNNC 1020
1993 ENERIKETEPDPSQEPKPAQSGTAPKSFHVEDTPVCSRRNSLSLSIDSEDLQEC 2042
1021 SPQSPHRAVD--VIGTRMSVSPAPPOQVNOEGHGRATGAYNTISRM-D-RHRVMD 1076
2043 ISSAMPKRRKPSRLKGDMKHSF-----RNM--GGLIGEDLTLDKIDQRPSEHGLSPD 2095
1077 HYSSDRDCEAD-----RO-----FYHRSRTE 1101
2096 SENFPMKALIOEGANIVSLHQAAAACLSRQASSSDSILSKSGISLGSFHLTPDQ 2155
1102 QPPL-----ERTTSSRSERPDN----- 1122
2156 EKPTSSNKGPRILKGEKSTLETTKIESKGIKGGKAYKSLITGKVNSEISQMKO 2215
1123 -LMRSMPLMTGRSAPSPALSRSHPRGVS-----QTSPTSPTGTRRGRQPLPLP 1174
2216 PLOANMPSISRGRTMHIIPGVNNSSTSPVSKGPPRLKTPASKSPSEG-----QTLATSP 2271
1175 KGTLESAAMDIEERRQKLNKYKQVAGSDPRLQDYHKKYSG--WDPHNGAD----- 1226
2272 RQAKPSVKSLSPPVARKQT-----SQIGGS-----KAPSRSGSDSPSPRAQOPLSR 2319
1227 TVSTKSSDSQDVAVASRTSSASRPSSTSYMSVQSERPRGNKISVFSKMONRQVGS 1286
2320 PLOSGRNSISGRNGISIPPNKLSQLPRTSSPTASTKSSGCKMS--YTS--PGRQM--S 2374
1287 GKNLTKSTISGDMCSLEKNDGSDQDTAVGALGTSKKRRSSIGAKMVAIVGLSRKSRSA 1346
2375 QONLTKQGLSKNASSIPRESGASGLQNMNGNANGAK-----VELSRSSSTK 2423
1347 SOLSTQEGGKKL-----RSTYQBSTETGLAVEMKMMWTROASREST----- 1388
2424 SSGSSDSRSEVRVILROSTFIKEAPSPTLRLKLEBSASFESLSPSSRPASPTPSQAQTPV 2483
1389 -----DCSMNYS 1397
2484 LSPSLPDMSLSTHSS 2498

RESULT 15
US-08-452-654-2
Sequence 2, Application US/08452654
Patent No. 5691454
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH

```

APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THALVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,654
FILING DATE: 25-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 08-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,035574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-452-654-2

Query Match 3.0%; Score 250.5; DB 1; Length 2843;
Best Local Similarity 18.0%; Pred. No. 1,1e-09;
Matches 305; Conservative 235; Mismatches 620; Indels 535; Gaps 70;

OY 30 SHUTEERKIIIAVMDROKKEEKEQSVLKIKIEHKAQPTQWPFSGITELVNNVLQPOQ 89
DB 1012 NIMDNDGGLDPIPNVSLKYSDEQ---LNSGRQSPQNERMARPKHILE--DEIKQSEQ 1065
OY 90 KQPKNEPQTKLHQQEKTKEDQYKKGESQO---QOEOKGDAPTGGICKTKFADGCGH 146
DB 1066 RQSRNQS-----TTPYVYTESTDDKHLKFPFHGOE-----CVSPYKSGANSET 1112
OY 147 NCSYQCTKFCARGGVLSRSNKKVMVWVNCILCRKQOELLITKSGAMFYNGSNTLQOPDKV 206
DB 1113 N-----KVGSNHGINQNVQSILC-----QEDDYEDDKPTVNSERYSEERQHE-- 1155
OY 207 PGLRNEEAPOEKKAKLHEQPOFOGAPGDLSPVAVEKGRAGHLTROPTINGSGVYKQIA 266
DB 1156 -----EERTNYSIKYNEKRHHVDPIIDYSL-----KYA 1184
OY 267 SDMPDRKRPSPVSRQNRNRYEEOSEEREDYSOYVSDGTMPRSPDYADRRSQREPOFE 326
DB 1185 TDIPSSQKQSFSSKSSSGGSKTE-----HMSSENTSTPSSNAKRONQLHPSSAQ 1237
OY 327 EPGLNLRDNRGRHRSKE-----YIVDDEDV-----ESRDEYERQREERYQ 370
DB 1238 SRGGQOKATCTKVSINQETIQTYCEDVPIPCFSRCSLSLSSAEDEIGCNOTTOEAD 1297

```

```

OY 371 A-----RYRSDPNLARIPIVKQPYEQEOMRTHAEVSRARHRRHSDVSLAN 415
DB 1298 SANTLQIAETKIGTKIGTSAEDPVSEVPVAVSOHPRTKSSRLQGS--SLSESRHNAVERPS 1356
OY 416 AELEDSRISLLMRP-----SPQSVSEERRAENORSYEMERTENO----- 459
DB 1357 GAKSPKSGAOTPKSPPEHYQETPLMFSRCTYS-----SLDSFSSRIASVQSEPCSG 1412
OY 460 -----GQSSYPORTSNHSPPTPKRSPILDLRPMRADSLRKHQHHLDPPSAVKRTKREK 513
DB 1413 MVSGIISPDLPDSPGCTMPPSRSKTTPP-----PPQTA--QTKRE-- 1451
OY 514 MEYMLRNDLSDDQSESVRPPPRPHKSKKGMQVSLSSSEEL-----ASTPEYTC 568
DB 1452 --VKNKAPTAEKRES-----GPKQAANAARVOVLPADLILLPATSTPDGSC 1502
OY 569 D-----DYEL-----ESESERKGDQ--KGRKT--SEOG 595
DB 1503 SSSLSLSLIDPEFIOKDYELRLMPVQENDNGNTESEQPKESNENQEKAKTIDSEKD 1562
OY 596 VLSDSN-----TRSEROKKMYTGGHSLLEDLEMSQPIKDSGYDTCS 639
DB 1563 LLDSDDDDIIELECIISAMPYTKSRKKGKPAQYASKLPPVARKPSOL-----PYK 1616
OY 640 TILNEHSHSDKHPTWQPSKDGRL--IGRIILN-----KRLKDGSVPRDSGAMIGL 690
DB 1617 LLPQNRLOPQKH-VSFYTGDDMPRYCYEGTPINSTATSLSDLTIESPPMLAGEGV 1675
OY 691 KVVGKMTESGRICAFITVKKGSGLADIVGHLRPGDEVLEWNGRLIQAGTFFEYVYIIIE 750
DB 1676 R--GG--AQSG-----EFKRDITPTG--RSTDEA-----QGGKTSSVTIPELD 1714
OY 751 SKPEQVELVY-----SRIGD-----IPRIPDSTHAOLESSSEFEOSKMDPSTISVT 799
DB 1715 DNKAEGDILAEICINSAMPKGSKHPRVAKKIMQYOQ--QASASSAPNNQOLDGKKKPT 1773
OY 800 SPMSPMLRLDVPQFLSGQSLIKIMLPDKVGHQLVTILGAKDLPRREDGRNRPYKATYL 859
DB 1774 SPVKP-----IPQNTERTV-----RKNADSKNMLMAERVF 1805
OY 860 PDRSDKNRRRTKTVKKTLEPK-----WNQFTIYSPVH-----RR 893
DB 1806 SDMKDSKQKMLKNNKSKDFNKLNNEDRVGSAFDPHHTPTLEGTPYCFSRNDISL 1865
OY 894 EFERMLEITLMDQARYRESEFLEIILELTALDLDEPHWYKLTQHDVSYLPLR-- 951
DB 1866 DFDDDVDILSR-EKAEILKAKENKESAKYTSHTELTSNGQSAKKQA--IAQPIRNGQ 1922
OY 952 PSPYLPRLQ-----HGESPTRIQOR-----SKRISQSVSDYCE-- 987
DB 1923 PKPILQOSTFFQSSKDIIPDGAATDEKLQNFALIENTPVCFSHNSLSLSDIDQENNK 1982
OY 988 -----DGVGVSDYDRHNG-----RDQSSTLSVEQYVMSNHC 1020
DB 1983 ENEDIKTEPPDSQGEPSKQASGYAPKSFVEVETPVCFSNNSLSLSDIDEDLLQEC 2042
OY 1021 SPQSGPHRVD--YIGTRSWSPSAPPQORVNEQCHRCSTRAGHYNTISRMQ--RARVMD 1076
DB 2043 ISSAMPKKKRSRLKQNEKISP-----RNM--GILGEDTLDLKDIDQPDSEGLSPD 2095
OY 1077 HYSSDRDRDCEAD-----RQ-----PYHRSRSTE 1101
DB 2096 SENFDWKAIQGANSIVSSLHQAANAACLSQASSDSDSILSLKSGISLGSFFHILTPOE 2155
OY 1102 QRPILL-----ERTTISRSESRPTIN----- 1122
DB 2156 EKPLTSNKGPRILKQGEKSTLETKIESESKIGKGGKVVYSLITGVYNSNEISGOMKQ 2215
OY 1123 -LMRSMPLMGRSAPPSPALSRSHPTGVS-----QTSFSSITPGGRKROLPOLR 1174
DB 2216 PLQAMPSPISNGRTMIHILPGVNNSSSTSPVSKGPKPLKTPASKPSPEG---QATTS 2271

```

```

QY 1175 KGTLEASAMDIERNRQMKLNKYQVAGSDPRLERODYHKKYRSQ---WDPHKGAD----- 1226
Db 2272 RGAKPSYKSELSPVARQT-----SQIGSS-----KAPSRSGSRDSTPSRPAQQLSR 2319
QY 1227 TVSTKSSDSDVSDYSAVSRITSSASRFSSTSYMVSQSERPRGNRKISVFTSKMÖNRÖGV 1286
Db 2320 PIGSPGRNISTPGRNGISPPNKLSQLPRTSSPTASTKSSGSGKMS-YTS--PGROM--S 2374
QY 1287 GKNLTKSTISGDMCSLEKNDGSDTAVGALGTSKKRRSSIGAKMVAIVGLSRKRSRA 1346
Db 2375 QÖNLTKÖTGLSKNASSIPRSESASKGLÖMÖNÖGANKK-----VELSRMSSTK 2423
QY 1347 SOLSQTEGGGKKL-----RSTVÖRSTETGLAVEMRNMTRQASREST----- 1388
Db 2424 SSGSESDRSRPPVLYRQSTFIKEAPPTLRKLEBASFEESLSPSRPASPTRQQAQTPV 2483
QY 1389 -----DGSNMSTSS 1397
Db 2484 LSPSLPDMSLSTHSS 2498

```

Search completed: November 21, 2001, 16:06:17
 Job time: 353 sec

THIS PAGE BLANK (USPTO)